西方数学文化理念传播译丛

丛书主编 汪 宇

Elementary Mathematics from an Advanced Standpoint

高观点下的初等数学

（第一卷） 算术 代数 分析

〔德〕菲利克斯·克莱因 著

舒湘芹 陈义章 杨钦樑 译

齐民友 审



复旦大学出版社

www.fudanpress.com.cn

内 容 提 要

菲利克斯·克莱因是19世纪末20世纪初世界最有影响力的数学学派——哥廷根学派的创始人，他不仅是伟大的数学家，也是现代国际数学教育的奠基人、杰出的数学史家和数学教育家，在数学界享有崇高的声誉和巨大的影响。

本书是克莱因根据自己在哥廷根大学多年为德国中学数学教师及在校学生开设的讲座所撰写的基础数学普及读物。该书反映了他对数学的许多观点，向人们生动地展示了一流大师的遗风，出版后被译成多种文字，是一部数学教育的不朽杰作，影响至今不衰。全书共分3卷。第一卷：算术、代数、 分析；第二卷：几何；第三卷：精确数学与近似数学。

克莱因认为函数为数学的"灵魂"，应该成为中学数学的"基石"，应该把算术、代数和几何方面的内容，通过几何的形式用以函数为中心的观念综合起来；强调要用近代数学的观点来改造传统的中学数学内容，主张加强函数和微积分的教学，改革和充实代数的内容，倡导 "高观点下的初等数学"意识。在克莱因看来，一个数学教师的职责是："应使学生了解数学并不是孤立的各门学问，而是一个有机的整体"； 基础数学的教师应该站在更高的视角（高等数学）来审视、理解初等数学问题，只有观点高了，事物才能显得明了而简单；一个称职的教师应当掌握或了解数学的各种概念、方法及其发展与完善的过程以及数学教育演化的经过。他认为"有关的每一个分支，原则上应看做是数学整体的代表"，"有许多初等数学的现象只有在非初等的理论结构内才能深刻地理解"。

本书对我国从事数学学习和数学教育的广大读者具有较好的启示作用，用本书译者之一，我国数学家、数学教育家吴大任先生的话来说，"所有对数学有一定了解的人都可以从中获得教益和启发"，此书"至今读来仍然感到十分亲切。这是因为，其内容主要是基础数学，其观点蕴含着真理……"。

第一卷目录

第一部分　算　术

---

① 此章中译本略去§1.4。——译者

第二部分 代 数

第三部分 分 析

博洽内容　独特风格

——《高观点下的初等数学》导读

吴大任

（一）书和作者简介

德国数学家F·克莱因的名著《高观点下的初等数学[①]》(以下简称《初等数学》)已由舒湘芹等同志译就，将由湖北教育出版社出版[②]。中译本出版，必将受到我国中青年教师和广大数学工作者的欢迎，对我国各级学校的数学教育也将产生巨大作用。

F·克莱因(1849—1925)是有深远影响的数学家。他的贡献遍及几何、代数、函数论、理论物理以及数学史等，在这些领域，他都留下了经典性著作。他是权威性的德国数学百科全书的主要创始人之一，曾任最高水平的德国数学年刊的主编。致力于这两项事业长达四十春秋。他热诚地献身于数学教育及其改革，是促进数学教育国际委员会创始人之一，并始终积极参与其中的活动。他著述《初等数学》这样的书，真可谓出色当行，游刃有余，得心应手。这书内容十分博洽，而论述生动活泼，不拘一格，把严谨性和直观性巧妙结合，深入浅出，使读者有举重若轻、左右逢源之感。

《初等数学》是克莱因的助手们根据他在哥廷根大学讲课内容整理而成的，分上下两卷。上卷“算术　代数　分析”(第三版)于1924年出版；下卷“几何”于1925年出版；英译本于1939年出版。60多

---

① “初等数学”指当时德国中小学的数学，比我国中小学数学略高。

② 本书曾由湖北教育出版社于20世纪90年代初出版过——编辑

年过去，数学面貌已有很大变化，我国目前的数学教育和德国当年也有很大差异。我们阅读这书时，对此必须注意。尽管如此，我们读来，对其内容和观点，仍然感到十分亲切。这是因为，其内容主要是基础数学，其观点蕴含着真理。当时德国数学教育中的不少问题，在今日我国仍然存在。克莱因声称，书是为中学教师和成熟的大学生写的，但按其内容，所有对数学有一定了解的人都可以从中获得教益或启发。

数学科学的整体性和数学教育的连续性

要想用一两句话来概括《初等数学》这本丰富多彩的书的特色，是困难的。也许可以说：它所展示的数学科学，是一个不断发展着的有机整体；克莱因所设计的数学教育，是一个随着数学发展而不断更新的连续过程。正如书名《高观点下的初等数学》所示，书的着眼点是初等数学，观点却是高等数学。数学各个分支，特别是数学两个基本对象——形与数结合起来了。讲算术、代数、分析时，总是充分运用丰富的几何图像。而讲几何时，用的是代数工具，又不乏几何语言。它还以大量篇幅阐述数学的各种概念和方法的发展与完善过程以及数学教育演化的经过。这些进程还在继续。

以下试对《初等数学》的若干具体特色作些介绍。

（二）《初等数学》若干特色

高观点

在《初等数学》的前言中，克莱因指出大学和中学数学教育的“双脱节”现象：大学生感到，他正在学的东西和中学学过的无关，而当他们到中学任教时，大学所学的用不上，因而那些内容就只存在于美好的记忆中。本书的直接目的自然是要改变这种不合理现象，以便把数学的新进展中所产生的新观念渗入中学数学教育中，按我们现在的说法，就是使数学教育“现代化”。

克莱因所采用的书名表明，他认为教师应具备较高的数学观点。理由是，观点越高，事物越显得简单。例如在实数域里不好理解的某些东西，从复数域的观点看，就清楚了；在欧氏空间里某些不好解释的现象，从射影空间的观点看，就有满意的说明。下面分别举两个具体的例。

克莱因指出，在中学，关于对数的传统讲法是有明显漏洞的。他建议把对数函数作为等角双曲线下的面积来引进，既简单又明确。他又指出，在复数域里，对数是多值函数，作为实函数的对数只是其中无数多个值之一。所以，在复数域里，对数函数的本质才看得清楚。我们的教师，无论是否愿意（或可能）采纳克莱因所建议的引进对数方式，有一点是肯定的：如果他了解作为复数的对数函数，当他讲实数时，就会心中有效，有可能弥补漏洞，至少当学生提出疑问时，他能正确回答，应付裕如。

通过变换群来阐明不同几何的本质及其相互关系，本是克莱因伟大创见之一。《初等数学》曾用了很大篇幅来论述欧氏几何、仿射几何和射影几何的关系。我认为，中学几何是欧氏几何，但也涉及图形的仿射性质（如三角形的重心）和射影性质（如三点共线）。如果教师能区别各种性质，在教学中自然是有利的。克莱因举了一个例来说明局限在欧氏空间就不好理解的现象：两个二阶曲面一般相交于一条四阶曲线，但两个球面（二阶曲面）一般只相交于一个（实的或虚的）圆（二阶曲线）。原来，从射影空间观点看，可以认为，两个球面还相交于“无穷远虚圆”，而两个圆在一起，恰好构成一条（退化的）四阶曲线。

教师应是多面手

克莱因对教师的要求是很高的。《初等数学》涉及的面很广。除正文 4 大部分外，还有两个附录：“关于 e 和 π 的超越性证明”[①]和“集

---

① e 是自然数的底，π 是圆周率。

合论”。每一大部分的写法和通常写法都很不相同，且其内容有不少超出通常写法的习惯范围。例如在“算术”部分写了四元数；在“几何”部分写了高维(以至无穷维)空间，并且随时讲到历史和应用。显然，克莱因认为，教师对这些都应当掌握或了解。他认为，大学生学到的具体东西不少，而许多重要的，以及在中学任教中用得着的东西却往往被忽视了。《初等数学》就着眼于弥补这些缺憾，揭示数学各部分之间的联系，指出它们的共性，它们产生与成长的内因、外因和过程以及它们的应用等等。克莱因认为，教师掌握的知识要比他所教的多得多，才能引导学生绕过悬岩，渡过险滩。他喜欢用“融合”这个词。《初等数学》也确实体现了初等数学同高等数学的融合，数学各部分的融合，几何观念和算术观念①的融合，感性与理性的融合(甚至一维、二维、三维空间的融合)等等。可以认为，全书是以上各种融合的融合。强调这一切，是为了使大学生和教师对数学有较全面的观点，有较高的修养。

数学发展的历史

克莱因反复强调的一个教育原则是按照学生的认知规律(包括年龄及成熟程度)进行教学。具体地说，要由简单到复杂，由低到高，由感性到理性等。他讲数学历史，是因为，他认为学生对数学的认识，在某个意义上，是人类对数学认知的历史过程相应的。当然，这绝不是说，学生的认知要重复历史上人类的认知。

在讲述数学的历史时，克莱因强调对事物认识深化的必然性(这不排除偶然性)。某些新概念的出现，是由于客观条件已经成熟而非产生不可。例如他指出，负数和复数的出现，是不以数学家的意志为转移的。非欧几何产生后，许多数学家是被迫承认它的。微积分由粗糙到严格，有着艰辛的历程。函数概念和几何对象范畴等的演化，都有过漫长的过程。我以为，了解一些历史是很有意义的；我们的课

---

① 在这里以及许多其它地方，“算术”是广义的，用来表示纯几何的对立面，包括代数和分析。

程往往分别构成首尾完整的逻辑体系。学生在学习中很难充分领会到数学是如何逐步成长起来,它又将如何继续发展。

公理体系

《初等数学》多处谈到公理体系,特别是关于数的公理和几何公理。克莱因认为,公理不能脱离直觉,不能排除人对客观事物的认识。因而反对那样一种观点:认为公理可以随心所欲地选取,只要它们彼此相容,即不产生矛盾就可以了。他还认为,不能按照公理体系进行教学。因为这首先不符合学生的认识规律。逻辑不是数学教学中的唯一指导思想。此外,他还有一个更深刻的理由。他把数学比作一棵树,公理比作树的根,当树逐渐长大时,躯干和枝叶向上长,同时根也向下长。因此既没有最后的终点,也没有最初的始点,即没有进行教学的绝对基础。至于教师,之所以要了解公理对数学的作用和意义,则是和他对教师的要求一致的;公理体系在数学作为一个演绎的逻辑结构中,毕竟占有极其重要的地位,不了解它就不能了解数学的本质和全貌。而在教学中,教师固然要考虑大多数学生的兴趣和接受能力,同时他又应能满足一些才华出众的学生的求知欲望,适当地回答他们可能提出的问题。

尺规作图和费尔马大定理

这两个问题在《初等数学》中并不占重要位置,但克莱因对它们的几句精辟议论,却可以用来作为对我们许多青少年学生和业余数学爱好者的忠告。《初等数学》较详细地讨论了用圆规和直尺作图问题。在谈到三等分角问题[①]时,克莱因指出:许多人拿出自己的“解法”,希望别人指出错误所在,但他们的知识基本上限于初等几何,又不肯去了解利用算术方法早已作出的不可能性证明。为了使读者对

---

① 即只用直尺和圆规把一个任意给定的角分为三等分的问题,这所谓“几何三大问题”之一,另外两个问题是“化圆为方”和“倍立方”。它们是古老的课题,但早已证明都是不可能的。化圆为方问题同圆周率的超越性有关,其他两问题之不可能是用算术方法来证明的。

这种算术证明有所了解,以便当他们接到送来的"解法"时,能站稳脚跟,他给出了用直尺和圆规不可能作正七边形的证明。

费尔马大定理最近几年才有了重大突破,但尚未最后解决。《初等数学》对这个"定理"的涵义有个十分有趣的图解,对它的历史直到克莱因时代的研究状况有简明的介绍。克莱因指出,自从 1907 年人们获悉解答这问题(即证明或否定费尔马大定理)的人会得到高额奖金后,就出现了大量的"证明"。这些人属于各行各业,但他们有一个共同点:完全不了解探索这个问题所遇到的严重数学困难,也不想去了解困难所在,只妄想靠突发的灵感就一下子加以解决。他们的结果当然是毫无意义的。

(三)对我们的启发

以上对《初等数学》的管窥蠡测,不求全面,但求无大错,可告无罪于该书作者和本文读者。下面结合我国现状,谈几点个人浅见,统请高明指教。

中学数学教师的提高方向

许多统计数字表明,我国中小学教师中有很大百分比没有达到教育领导部门所规定的最低业务标准。这里不谈这些现象存在的根本原因(如教育投资长期太少,教师待遇过低),只谈教师提高的方向。我以为,拓广教师的知识领域,提高他们的教学修养,是当务之急。为此,一个非常重要的策略是,必须把教师从"题海"中解脱出来。不少教师抱怨,经常要花大量时间和精力去收集习题,把解题方法分类,编写习题解答等等,根本顾不上进修。而不那样做,四面八方又不谅解。教师的这种苦衷,了解的人恐怕不多。事实上,"题海战术"对广大学生也是利少弊多。用各种方式帮助现有中小学教师提高,高等学校有责任,也有余力。在中小学教师大半已达到规定标准后,这个标准还应有所提高。

初等数学教育现代化问题

若干年前，许多国家进行过数学教育现代化的研究和试验。现在谈论它的人似乎少些了。我以为，问题不在于要不要现代化，而在于如何现代化。有一条原则是必须坚持的，即要按学生的认识规律进行教学。用现代数学知识武装中学教师，是初等数学教育现代化的前提。

大学数学系的任务

师范院校要面向中学的原则已经定下来了，“向综合大学看齐”的倾向也已经改过来。其实，我以为，师范院校只要注意保持“师范”特色，综合大学数学系的课，师范院校也可以开设。我说的是“可以”，不是“全部必须”。因为中学教师掌握这些课的内容有好处。至于综合大学数学系毕业生也可以（甚至必须）有一定比例到中学任教。那种认为综合大学毕业生到中学教书是“大材小用”的说法，是站不住脚的。为什么大学生和研究生报名当旅馆服务员就不算“大材小用”？在许多国家，师范院校以外的大学毕业生还要通过教育课程的考试才能取得中学教师的资格呢。可见问题的根本在于教师的待遇。

大学数学教育的改革

大学数学教育也大有改革余地。例如必修课分量偏重，“上层建筑”要求偏高，基础不全不牢等等，都不利于人才的健康成长；在大量招收研究生后更是如此。在这里，我只着重谈谈几何形象问题。许多数学大师都强调形与数的统一。希尔伯特说过：“算术记号是写下来的图形，几何图形是画下来的公式”。克莱因认为：几何基础可能要以算术为起点，却不能脱离几何直观，而且他讲算术问题时，总要结合几何图像。他们的观点是完全一致的。问题是，在我们的高等数学教育中，几何形象被严重忽视了。作为基础课的解析几何已不能保持最低限度的分量。许多代数和分析课强调自我演绎体系，从逻辑和审美观点看很好，缺点是形与数固有的内在联系割断了。纯几何的演绎体系似乎已逐渐成为历史，为几何、算术、代数所取代，但

也不能因此而抛弃几何直觉。另一个问题是，我们很少对学生介绍数学发展的历史。在这两方面，我认为综合大学有不少地方可以向师范院校学习。我们并不需要在综合大学数学专业恢复 50 年代作为必修课的几何基础和数学史，但可以通过改革教学内容和方法来达到加强几何形象的目的。当然，这涉及教师的培养与提高问题。

善于数学的"热门课题"

在我国青少年学生和业余数学爱好者中，"几何三大问题"（主要是三等分角问题）和费尔马大定理（以及哥德巴赫问题）都是（或曾经是）"热门课题"，但他们"研究"的质量似乎比克莱因时代的德国还低。其实前者已证明为不可能，后者即使在数学界也只是"热门话题"而不是"热门课题"。它们在我国某些人中之所以成为"热门"，部分原因是他们片面理解"解放思想"，更重要的是我们宣传教育不够，我们希望教师们能做这些人的工作。对于执着要搞这两类问题的人，在肯定其精神可嘉之余，要教育他们尊重科学，实事求是，适当地向他们"泼冷水"；鼓励学生打好基础，鼓励业余数学爱好者把精力和时间用于更能发挥自己专长的地方。

一点希望

希望我国有众多人像克莱因那样关心数学教师的培养与提高，关心数学教育改革，并为此做些实事。《初等数学》中译版的现实意义就在于，它将促进这两方面工作的进程。但是德文本出版已过了 64 年，英译本出版也过了 50 年。现代数学已发生了极大变化，新成果、新概念、新观点、新学科层出不穷。但是数学的本质与真理是永恒的，像克莱因那样探索数学教育的规律，当是一以贯之的。我们热切希望我国高水平的数学多面手会写出更结合我国实际的、现代化的《高观点下的初等数学》。这样一本书出版，将是我国数学教育史上一件大事。

1989 年 6 月于南开大学

纪念克莱因

——介绍《高观点下的初等数学》

齐民友

我们不妨用这部名著最后一卷(第三卷)的最后一句话来开始我们的介绍。

"保持一流大师的遗风:回到固有的生动活泼的思考,回到自然!"

这段话栩栩如生地刻画了本书作者克莱因的风范。他是近几个世纪的当之无愧的数学大师。在解释这一段话之前,我们先来介绍一下他的生平和对于数学事业的贡献。

克莱因(1849—1925年)生于德国莱因河畔的杜塞尔多夫。中学毕业后进入波恩大学,师从普吕克(J. Plücker)。当时,普吕克的科学兴趣集中在几何学,所以克莱因也以几何学开始了自己的数学生涯。克莱因得到博士学位时,恰好普吕克去世。他也就离开了波恩。在好几所大学工作以后,他受到著名几何学家克莱布什(R. F. A. Clebsch)的青睐,得到了埃尔朗根大学的教职。1872年发表了著名的就职演说,题为"Vergleichende betrachtungen über neuere geometrische forsuchungen"("近代几何研究的比较评论",英文本 Felix Klein, *A comparative review of recent researches in geometry*),也就是著名的爱尔朗根纲领。遗憾的是,英文本不太容易找到,所以尽管它影响深远,读过它的人却很少(作者遗憾地承认,自己也没有读过。过去可以说是因为很难找到,但是现在可以在网上找到全文:Math. ucr. edu/baez/Erlangen/Erlangen_tex. pdf)。当时克莱因还只有23岁。尽管他在几何学上如此贡献卓著(另一项贡献是给出了双曲几何的克莱因模型,

并且证明了,这种几何学的相容性等价于欧氏几何的相容性),他自己却认为自己在数学上最大的贡献是在复分析。他认为自己最大的成功在于发展了黎曼(B. Riemann)关于解析函数理论的几何物理的途径,把它与群论、不变式理论、高维几何学、微分方程等融合在一起。1880 年起他来到莱比锡大学,而且树立了一个目标,就是按照黎曼的思想建立一个学派。可是天忌英才,从 1882 年起,他就因重病而不能继续从事这项伟业。1886 年,他离开了莱比锡去哥廷根,可以说,他的研究生涯至此结束。

有人说,克莱因有两个灵魂:一方面他渴望宁静的研究生活;另一方面他又是热情的教育家、组织者。从他 1886 年到哥廷根以后,他就致力于把哥廷根建成为当时一流的数学中心。希尔伯特(D. Hilbert)就是克莱因延聘到哥廷根的。著名的刊物 *Mathematische Annalen*(《数学年刊》)也在他的主持下,成了当时最有权威的数学刊物之一。而十分重要的是他对于中学数学教育改革的贡献。他在哥廷根一直为中学教师讲课,讲稿最终整理成《高观点下的初等数学》这部名著(以下简称《初等数学》)。从 1901 年手稿面世直到 1928 年第三卷由赛法特整理成书,历时 27 年。其时克莱因已经去世 3 年了。

已故的吴大任教授,为中文译本写的序言对这本书作了非常精到的评述。现在在数学界,凡说到初等数学,不少人心中总会有一种"小儿科"的感觉,吴先生指出《初等数学》就是基础数学。而且这里讲的不是什么搞博士点、搞重点学科那种意义上的基础数学,而是整个数学的基础。本书原名的 Elementarmathematik 是否也可以如是理解呢? 如果是这样,那么本书的读者,就不只是中学老师,而用吴先生的话来说:"所有对数学有一定了解的人都可以从中获得教益和启发。"也正因为如此,"《初等数学》一书,至今读来仍然感到十分亲切。这是因为,其内容主要是基础数学,其观点蕴含着真理……"为什么要从高观点来看呢? 吴先生说:"理由是,观点越高,事物越显得简单。例如在实数域里不好理解的某些东西,从复数域的观点来看,就清楚了;在

欧氏空间里某些不好解释的现象，从射影空间的观点来看，就有满意的说明。”他接着说：“克莱因认为，数学专业大学生学到的专业知识不少，而许多重要的……却往往被忽视了。《初等数学》正是着眼于弥补这些缺憾，揭示数学各部分之间的联系，指出它们的共性，它们产生与成长的内因、外因和过程以及它们的应用等等。”吴先生接着强调，《初等数学》特别着重“融合”，即“初等数学同高等数学的融合、数学各部分的融合、几何观念同算术观念的融合、感性与理性的融合等等。可以认为全书是以上各种融合的融合。”这里，还对作者原意给予补充：数学与物理学以及各种自然科学的融合，数学的逻辑结构与历史发展的融合。

现在对于克莱因关于中学数学教育改革的具体主张作一些介绍。这些主张中的要点就是把微积分初步放在中学，而且强调函数的概念。请注意，这是 100 年前提出的，而我们不少人却认为是新一轮课改的成果(或隐患)。放眼世界，在大多数发达国家，这已经成了不需要讨论的事情(可能美国例外)。有许多国家写出了很好的教材。我愿特别呼吁：请注意日本的中学教材。我曾比较仔细地读了日本数学家藤田宏主编的高中(日本人称为“高校”)数学教材。据说日本的义务教育是 12 年，即相当于我国的高中。又据说，日本高中毕业生有 40%学理科教材。也就是说，基本上掌握了一元函数的微积分。如果我国每年进入高校的新生有 20%(按今年大学招生数估计，即 120 万人)达到这样的教材所反映的数学水平，那么，我国许多高校理工科的数学教学，将面临新的挑战。美国数学会曾翻译了小平邦彦主编的高中教材，受到不少好评。例如 *Zentralblatt MATH* 在其评论中就说，应该感谢美国数学会，组织翻译了一本比现在欧美国家使用得更好的数学教材。是否有哪一家出版社愿意考虑这件事？回到本题，为什么要在中学教微积分初步？克莱因说，如果没有这样的准备，就不可能理解当前正在研究的自然现象。这与另一批美国学者的看法是一样的。美国德州仪器公司(TI)一位高官组织了一些数学家研究美国的数学

教育，他们发现，尽管美国数学界和数学教育界看法有分歧，但在一点上却有共识：21 世纪的劳动者应该懂得微积分初步。可是克莱因提出这个观点早在 100 年前！

关于在中学教微积分初步，最大的疑虑来自对中国现在的中学教学水平的估计。确实，根据我近年的了解，相当大的一批中学，如果能够把一本比较平实的教材，例如几年前用的老教材教清楚，就很不错了（还要花很大的力量）。不平衡仍然是基本的国情。但是如果只考虑这一部分学校，显然也是不行的。如何对待这种不平衡性，当然是另外一个问题了。

同样，从我接触到的许多中学来看，教师的数学水平经过一定的努力，是能够满足现行"课标"的要求的。而正是对于这样一大批中学（和大学）数学教师，克莱因在《初等数学》一书中提出的意见更是值得深思。《初等数学》第一卷第四章后有一个附录："关于数学的现代发展及一般结构"，克莱因对自 17 世纪以来围绕微积分的数学发展的轨迹作了十分精彩的论述。概括地说，克莱因指出，这里有 3 种不同的进程，互相交替，又互相补充。片面地只强调其一，而忽略其他，是有害的。所谓进程 A，在教学上的表现，就是：

（1）先是方程和有理式的形式运算，用根式解方程；

（2）系统地研究幂运算及其逆，出现了对数；

（3）现在进入几何领域，有了三角函数，然后三角成了一门独立的学科（或章节）；

（4）进入"代数分析"，也就是以幂级数为中心，讲二项式定理、指数与对数函数的展开式，$\sin x$，$\cos x$ 的展开式。这时突然出现了欧拉（Euler）公式 $e^{ix}=\cos x+i\sin x$。学生难免奇怪，何以来自全然不同领域的函数会有如此奇怪的关系！

（5）这时进入了复域，就是以幂级数为中心的魏尔斯特拉斯（Weierstrass）理论。

很容易看到，我国的数学教学（从高中到大学的分析数学各课程）

就是按照进程 A 来组织的:(1)到(3)是高中教材,(4)和(5)是大学教材。如果说有区别的话,就是在(3)与(4)之间,插进了一块或大或小的极限理论。特别是大学数学系本科,是完全以 $\varepsilon-\delta$ 为基础的魏尔斯特拉斯理论,使学生吓得要命,所以人们戏称为“大头微积分”。不但是头重脚轻,甚至是头足倒立。因为它不由分说地把历史发展的“终结”(魏尔斯特拉斯是在 1861 年在柏林讲课时才使用了这种讲法)放在历史的“起点”(牛顿(J. Newton)和其他人在 1650 年以前就在用幂级数了,这就是(4)的内容),提前了大约 200 年!那么进程 A 有什么指导思想呢?克莱因指出,“进程 A 是建立在把一门学科进行分解的概念上的,即把一个整体分成一系列互相独立的部分,使各部分独立发展。尽量少借助于其他部分的知识,尽可能避免引入相邻领域的概念。进程 A 的理想是把各个局部领域的知识结晶为一个逻辑封闭系统。”

那么,在数学的发展中还有没有别的进程呢?克莱因指出,还有进程 B。克莱因指出它的中心思想是解析几何的思想,就是我们常说的“数形结合”。具体来说就是:

(1) 先从最简单的函数:多项式,有理函数的图像开始,得出一个概念:函数的零点(即方程的根)就是图像与 x 轴的交点。

(2) 有了图像就自然地出现了斜率、面积的问题,于是微积分出现了。

(3) 许多(甚至是绝大多数)函数的积分不能用已知函数来表现。例如由双曲线 $xy=1$ 下面的面积,得出了对数函数的定义

$$\int_1^x \frac{\mathrm{d}x}{x}=\log x,$$

(请注意,这不是克莱因别出心裁想出来的讲法,历史就是这样的,牛顿很明确地这样做,纳皮尔(Napier)也是用积分或微分方程来定义对数的)。类似于此,研究圆扇形的面积给出了反三角函数:$\int_0^x \frac{\mathrm{d}x}{\sqrt{1-x^2}}=\arcsin x$。克莱因指出像这样走下去,就会得到椭圆函数

和椭圆积分,而它们确实是克莱因时代数学的高峰。

(4) 通过一个统一的原理:泰勒(Taylor)级数,得出许多函数的幂级数展开式。在历史上牛顿确实是从上面的两个积分得到 $\log(1+x)$ 和 $\arcsin x$ 的级数展开式,然后又十分巧妙地对这些级数进行反演(即求反函数的展开式)得到了 e^x 和 $\sin x$ 的展开式。请注意,这是历史的真相!

(5) 循此前进,在复域中得到黎曼的几何化的复分析。

回到我国的数学教学,大概(1)和(2)还是大家熟悉的,以后就渐行渐远,可能很多人还会以为以后的几条多是小玩意,不必去关心了。克莱因则说,"进程 B 把主要重点放在各局部领域的有机结合上,放在各个局部的互相促进上,因而宁可采用统一的观点来理解好几个领域的方法。进程 B 的支持者的理想是把数学科学的总和理解为一个巨大的相互联系的整体。"

除此之外,克莱因还指出,数学的发展还有进程 C,其特征是强调算法的作用。不过这里的算法一词和我们理解的不太一样,而更加广泛,大体上是指包括用字母和各种符号进行演算和推理的过程。《初等数学》一书对此没有展开,我们也就不再讨论了。

人们会问,这 3 个进程,孰优孰劣?其实克莱因没有这样提出问题。存在 3 种进程,是历史的真实,所以他认真详细地讨论了历史,主要是以人们认识各种函数的历史,说明这 3 种进程如何互相作用,彼此消长,而终于进程 A 占了上风。但是克莱因很明确地指出,从教学的角度来看,这有明显的弊端。他说,对于一个数学家和数学教师,倾向于,或者更喜欢哪一种进程,无所谓是非好坏,但是对于广大学生,克莱因十分明确地认为,现在进程 A 的地位是太高了。"人们不免要问这两个方法(即进程)哪一个更有效?对于没有特殊的数学抽象天赋的学生,哪一种方法更好?"对于我们这些教书的人来说,这真是一个好问题。我们都有这样的经验:每一个班上总有几个好学生,他们中的许多人确实有特殊的数学抽象天赋,在 $\varepsilon-\delta$ 的海洋中可以说是如

鱼得水。但是多数人则不一定如此。再说，这少数人能否真正再上一层楼，还需要努力。看来问题的关键是缺少进程 B 的熏陶。克莱因以指数函数和三角函数为例来说明这一点。按照进程 A 我们用

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots$$

作为指数函数的定义。在经过微分学很长的旅行以后偶然地(即与学生熟悉的几何意义完全无关地)发现

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!},\ \sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}。$$

最后把 e^x 中的 x 换成 ix，又完全偶然地得到欧拉公式。为了强调这种做法的意义重大，特别声明这个毫不自然的做法给出的竟然是自然对数的底。在我国许多大学和中学教材中还要特别声明 $e = 2.71828\cdots$ 是一个无理数(为什么不说是超越数呢?)，可能想让学生联想到 π。可是在什么地方出现了圆周呢？另一个与此相关的例子是在正态分布中还出现了 $\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$，怎么人的生老病死又扯了进来呢？大概教师们都会说 e 实在太美了、太神奇了，欧拉公式是最美妙的公式等等。学生是不是也这样想，就是另外一回事了。克莱因这本书写得很客气，因为他绝无轻视进程 A 以及喜欢这个讲法的学者们的意思(如果读者不信，请读一下第三卷关于处处连续但不可微函数的介绍，关于皮亚诺曲线的介绍等等，请与您所熟悉的数学分析教材比较，看看是谁讲得更清楚)。但是在这里，克莱因讲了几句重话：$e = \lim\limits_{n\to\infty}\left(1 + \frac{1}{n}\right)^n = 2.71828\cdots$，“这个定义，通常都放在大部头的分析教科书最开始处，这是模仿法国的教法。而丝毫不讲它的来由，这样就缺少了真正最有价值的、能促进理解的部分，即不解释为什么恰好用这样特别的极限做底，为什么由此导出的对数称为自然对数。”然后，同样形式地给出了自然对数函数的展

开式 $\log(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\cdots$，完全不讲对数表是怎样得出的。(当然，现在我们都不用对数表了，计算尺对于现在的学生早就成了古董。但是数值计算难道会成古董吗?)克莱因对此大为不满，他斥责说："这种不问究竟的态度是可鄙的实用主义，它藐视任何一种比较高级的教学原则，对它必须严厉谴责。"大家知道，克莱因是计算数学和应用数学作为专门学科的开创者之一。《初等数学》一书对于现在我们认为应该放在计算数学课程里的内容：插值法，包括三角插值、数值积分等等，无不用充分的篇幅加以讲解，而对轻视者一定严厉谴责。这与我们现行的大学教材(包括工科教材)形成了鲜明对照(这一段中引用克莱因的话请参看第一卷第七章 § 7.1)。

再来看看按照克莱因的想法，按照进程 B，指数函数及其与三角函数的关系应该怎样教。他说(以下的引语又见第一卷第四章的附录)："在进程 B 中，其中联系是以很容易理解的方式出现的，并与这些函数的意义相一致，这一点是从一开始就加以强调的。事实上，函数 e^x 和 $\sin x$ 在这里是同出一源的，是从求简单曲线的面积产生的，由此很快就把人们引到最简单的微分方程

$$\frac{\mathrm{d}e^x}{\mathrm{d}x}=e^x,\ \frac{\mathrm{d}^2\sin x}{\mathrm{d}x^2}=-\sin x,$$

这两个相应的微分方程当然是上述一切应用的基础。"

我们就不再往下引用了。现在我们常说什么什么"理念"，克莱因当然是有丰富理念的人。但是更值得我们学习的是，克莱因尽管有崇高的地位，可是他没有停止在理念上或将具体事情让别人(例如学生)去做。在《初等数学》这本大书里，他可以说是一个一个函数地研究，讲它们的历史、应用，提出教学建议。"理念"溶化在数学中，呼之欲出，而又不显踪影。这样，读这本书，您会感到极有收获，而不得不心悦诚服。不得不承认克莱因是真正的大师！克莱因没有叫我们去做什么，但是你会自己也想仿效他去做。可是这绝非轻松事！上

面讲的欧拉公式等等，听起来确有道理，可是想把它变成大学生（甚至中学生）也能懂的材料，恐怕要费极大的工夫了。我愿向读者推荐一本书：T. Needham, *Visual Complex Analysis*，人民邮电出版社出版了它的影印版，希望不久之后可以看到中文译本。看看人家是怎样把它变成大学生（可能还有优秀的中学生）能读的书。可是还有进一步展开的空间。有条件的中学老师也可以读一下它的第一章。这样就会知道，所谓爱尔朗根纲领是怎么一回事了。

由于本书内容丰富，我们只能选取一段来介绍克莱因的思想。在结束本文时，我愿再回到吴大任先生的序言。吴先生是我国数学界受人尊敬的前辈和长者。本书第三卷由他和夫人陈鹀亲笔翻译，是他们夫妇对我国数学界和青年人最后的奉献。当时吴先生视力下降，可是在这样的困难情况下，仍然字迹工整，十分令人感动。他在中译本序言的最后表达了自己的殷切希望："（本书）德文本出版已过了 63 年，英译本出版也过了 49 年，现代数学已发生了极大变化，新成果、新概念、新观点、新学科层出不穷。我热切希望我国高水平的数学多面手会写出更结合我国实际的、现代化的《高观点下的初等数学》。这样一本书出版将是我国数学教育史上的一件大事。"

再读一下克莱因的话：

"保持一流大师的遗风：回到固有的生动活泼的思考，回到自然！"

这本书不正是生动地向我们展示了什么是一流大师的遗风吗？

2007 年 6 月 24 日于武汉大学

第一版序

我向数学界，特别是向中学数学教师奉献的这一卷新书，应该看做是"中学数学教学讲义"，尤其要当作为去年由杜伯纳(Teubner)出版社出版的、席马克与我合著的《数学教学组织》一书的续篇。那时我们所关心的是向数学家介绍种种不同的教学方法，而目前我所关心的是数学教学内容的种种进展。我的努力方向是充分结合当今的数学教学方法，从现代数学科学观点出发，向数学教师以及已成熟的学生介绍数学教学的内容及基础，但要尽可能做到简洁、有启发性，也有说服力。我不准备仿照韦伯(Weber)和韦尔斯坦因(Wellstein)两人那种有系统的讲法，只是想像实际讲课中那样，随机应变、自由发挥。

这样的写作大纲，在我和席马克去年4月出版的那本书的序言中已有提及，目前仅仅贯彻于本卷算术、代数、分析这3个部分中。我曾希望，尽管有许多困难，席马克先生会挤出时间把我的讲义整理好再付印。不过怪我不好，不断要求他把时间用在我们两人都感兴趣的其他教学法问题上。很快就发现，当初的计划不能完成，特别是不能在短期内完成。可是要想对刚刚浮现的那些教学问题发挥真正的影响，非要在短期内完成不可。于是像前几年那样，只得求助于一种比较简便的方法，将我的讲义付之石印，尤其是因为我现在的助手——欧内斯特·海林格(Hellinger)博士特别胜任这个工作。我们不应当低估海林格博士付出的劳动。把一个教师在偶然条件影响下讲的话整理成通顺的记录，这不是一件轻而易举的事。通常要印刷出版就要做到叙述精确、解释一致，可是讲课的人往往做不到这

一步。

对于是否继续出版有关数学教学的著作，至少是在几何方面，我很犹豫。我仅希望这一卷石印的书能促使高中数学教师再度运用独立思考，确定讲授教学材料的最好方法。这本书的目的只不过是起到智力启发作用，而不能当作一本详尽的手册。至于手册之类，应该让积极从事中学教学的老师来写。某些人可能误会，以为我的书还有别的什么目的，这样想就错了。特别应说明，德国自然科学家及医师协会教学委员会拟的课程大纲不是我写的，只是在我的协作下由一批优秀的中学数学教师写成的。

最后，关于本书的叙述方法，我只想说明一下，像以往一样，我在这里做出的努力，就是把几何直觉与算术公式的精确性结合起来，而我感到特别高兴的是能遵循各种理论的历史发展，理解今天教学上互相平行的几种教授方法的显著差别。

克莱因(F. Klein)

1908 年 6 月于哥廷根

第 三 版 序

自从施普林格(Springer)公司令人钦佩地出版了我的科学著作集以后,经柯朗(R. Courant)教授的建议,该公司又提出以书本的形式出版我的讲义,因为这些讲义本来是 1890 年以后以石印印刷形式出版的。除了少量存本外,现已售缺。

以前由杜伯纳出版社接手发行的这几卷讲义,主要是以往几十年中我的各个助手的记录稿。一开始我就知道,如果不找年轻人帮忙,我是无力担负起修改的任务的。很久以前,我就表示过这样一个信念,人超过一定的年龄,就不该发表独立的作品。上了年纪的人也许有资格对一本新书的写作进行一般的指导,但已无能力详细整理材料,把文献中的新进展都考虑进去。因此,在保证向我提供充分协助之后,我才接受了施普林格出版公司的提议。

这几卷曾以石印刊行的讲义分为两类。较早写的一部分是我不断讲授的专门的讲义,写这些讲义,纯粹是为了使下一学期的学生手头有我已经研究过并建议作为进一步研究基础的材料。这包括非欧几何、高等几何、超几何函数、线性微分方程、黎曼面及数论讲义。与此对比,我还出版了几本本来就准备给广大读者看的石印讲义,即:

一)册是《微积分在几何中的应用》,这是 C・H・缪勒(C. H. Müler)根据他的笔记整理的,该书的目的是为了填补应用数学的需要与纯数学的新近研究成果之间的空白。

二)册和三)册是《高观点下的初等数学》两册,这是 E・海林格根据他的笔记整理的,目的在于促使中学数学及理科教师了解他们在大学里学的东西,特别是纯数学研究对于他们教学的重要性。

第二类讲义看来不必作彻底的修改,有的地方作若干润色,同时补充一些注解就可以了。值此出版之际,采取了初步的加工。二)册、三)册、一)册(按此顺序)合成一大册,内分一卷、二卷、三卷,冠以《高观点下的初等数学》的书名。这样把一)册与二)册、三)册合并,理解应用数学在现代中学教育中日益重要的教师,是会欢迎的。

同时,第一类讲义的修改也已从《非欧几何》一书开始。但是要把那本书整理得严谨而全面,并照顾到最新的科学进展,非要做比较彻底的修改不可。我的大致计划就是如此。现在就本书的第一卷说几句话。

我重印了本卷 1908 年版的序,是因为它讲清了这一卷的产生经过①。同样石印刊印的第二版(1911 年),没有作重大的修改,次要的附注现已收进正文,不另提及。本版基本上保留了第一版的正文②,但含有成书当时的特点,否则就必须改变全书的结构,失去了一致性。不过自第一版出版以来的 16 年中,科学已有进展,我们的中学教学体系也发生了很大的变化,而且还在变化之中。这种情况在塞法特(Seyfarth)博士和我合写的附录中已经考虑到了。弗特博士还对本书作了必要的文字风格上的修改,照料了印刷事宜(包括插图),所以我对他表示真诚的感谢。我以前的几个同事——海林格先生、弗尔迈尔(Vermeil)先生以及哥廷根的A·瓦尔特(Walther)先生,在校读清样时都提出了许多有益的建议。特别要感谢弗迈尔及比利希先生编了人名表及索引(中译本中已删去——译者)。出版家朱利叶斯·施普林格在极其困难的条件下出版我的这本书,又一次说明了他刊行数学著作的决心名不虚传。

克莱因

1924 年复活节于哥廷根

---

① 这里提到的我的同事 R·席马克(R. Schimmack),他在伏案工作时因心脏病突然发作,已于 1912 年逝世,卒年 31 岁。

② 新的评注放在括弧内。

英文版序

菲利克斯·克莱因(Felix Christian Klein)教授是德国有名的数学研究家,他也是一位循循善诱的教师。他以罕见的天才,集一切数学领域的知识于一身,并善于领悟这一切领域之间的相互关系。他认为使学生了解数学并不是孤立的各门学问,而是一个有机的整体,是他作为一个教师的明显职责。他对中学数学教学有浓厚的兴趣,不仅关心应该教些什么内容,而且关心怎样教才是最有效的方法。多年来在哥廷根大学为德国中学数学教师及在校学生开设讲座,已成了他的习惯。他一贯努力缩短中学和大学之间的差距,从传统的漠不关心中激起中学教师对高等数学的兴趣,把中学数学教学引向健康发展的方向;同时也努力扭转大学的态度及教学方向,使之承认中学的正常地位,使数学教育前后一贯。

克莱因教授的这些讲义最终印成3卷,名为《高观点下的初等数学》。这是一本无比珍贵的著作,同样可作为大学教师和中学教师的参考书。无论就材料安排的巧妙或就讨论方式的引人入胜来说,目前都没有一本书可以同这本书相比。本卷英文版是上述著作的第一卷。它的出版是哥廷根大学柯朗教授建议的结果。出版的目的是满足英语国家现任数学教师及未来数学教师的需要。我们诚恳地希望,我们的译作虽然有点意译,但是保持了原著的精神。

英文版译者

洛杉矶加利福尼亚大学数学教授

E·R·赫德里克(E.R. Hedrick)

伯克利加利福尼亚大学数学教授

C·A·诺布尔(C.A. Noble)

前　言

近年来[1],在大学数学教师及其他理科教师中,对如何更好地培养未来中学师资产生了广泛的兴趣。这确实是一种新的现象。在此之前,长期以来,大学里的人只关心他们的科学本身,从来不想一想中学的要求,甚至不考虑同中学数学的衔接。结果如何?新的大学生一入学就发现,他面对的问题好像同中学里学过的东西一点也没有联系似的。当然他很快就完全忘了中学学的东西。但是毕业以后当了教师,他们又突然发现,要他们按老师的教法来教传统的初等数学。由于缺乏指导,他们很难辨明当前教学内容和所受大学数学训练之间的联系,于是很快就坠入相沿成习的教学方法,而他们所受的大学训练至多成为一种愉快的回忆,对他们的教学毫无影响。

现在的改革运动就是要克服这种对中学教学和大学教学都没有帮助的双重的不连贯性。一方面要努力在中学教材中注入由现代数学进展得来的与现代文化相一致的新观念(本书中将不断有机会探讨这一点);另一方面又试图在大学教育中把中学教师的需要考虑进去。现在为你们开设的这样一个综合的课程,我认为正是帮助你们提高的最重要的方式之一。我所针对的学生绝不是初学者。我预计你们都已经熟悉数学各主要领域的要点。我时常要讨论代数、数论、函数论等问题,但不得不放弃细节。因此,要跟上我的思路,你们必须对这些领域有相当的了解。我的始终如一的任务是向你们指明一般课程中没有充分指明的各个数学领域中种种问题的相互联系,尤

---

① 再请注意:本卷文字几乎与1908年石印本完全相同,以后几年加进去的内容已收入附录。

其是强调这些问题与中学数学问题的关系。我希望通过这种方式使你们更易于掌握从大量放在你们面前的知识中汲取促进教学的养料的能力。而你们进行学术研究的真正目标,我认为就在于掌握这种能力。

现在请让我向你们提出几份近期的文件,其中包含许多宝贵的材料,说明对教师的培养已产生了广泛的兴趣。首先我想到 1907 年 9 月德累斯顿自然科学家会议上的发言,我们德国自然科学家及医师协会教育委员会向大会提出了关于未来数学和理科教师科学培养的建议。你们可以在这个委员会的全文报告①的最后一节找到这些建议。自 1904 年以来,这个教育委员会一直在讨论有关数学等自然科学教育的全部问题,现已结束活动。我鼓励你们不仅注意这些建议,而且注意这个很有趣的报告的其余部分。在德累斯顿会议结束以后不久,9 月 25 日于塞尔召开的德国语文学家及中学语文教师会议上,发生了同样的辩论。当然,是把数学教学改革当作语文学界平行改革的一环进行了讨论。在我作了有关数学等自然科学教学改革的目的报告之后,P·温德兰(P. Wendland)(布列斯劳)就考古学问题作了发言,亚历山大·布兰德尔(Al. Brandl)(柏林)作了现代语言的发言,最后由爱德勒·哈奈克(Ad. Harnack)(柏林)作了历史与宗教的发言。上述 4 个发言一起收入一本小册子②,我特别建议你们读一读。我希望,这个引人注目的开端会发展成我们自然科学家与语文学家之间的进一步合作,创造友好合作及相互理解的精神,因为过去双方关系虽然不能说是敌对,也至少不够协调。我们要永远培植这样良好的关系,即使我们有时在内部偶尔说他们一两句不好听的话,他们也可能说我们几句不好听的话。要记住,你们以后要到中

---

① *Die Tätigkeit der Unterrichtskommission der Gesellschaft deotscher Naturforscher und Ärzte*, A·格茨默(A. Gutzmer)编,莱比锡及柏林,1908 年。

② *Universität und Schule*, P·温德兰、Al·布兰德尔、Ad·哈奈克的发言,莱比锡,1907 年。

学里与语文教师共事，这就需要相互理解与合作。

除了我们数学界以外的上述改革努力，我还想提到几本书，这几本书同样以数学教学为方向，对我这些讲座极为重要。3 年前，我第一次抱着类似的目的开了一次讲座。我那时的助手席马克收集了材料，其中第一部分已出版[①]。那本书里考虑到了包括大学在内的各种类型学校以及各类学校的数学教学和共同兴趣等之类的问题。在下面的讲座中我会时时提到上述书中的内容，但不加以重复，以便深化这些讨论。那本书所讨论的，是学校中数学教学的安排，现在要讨论的是数学教学的内容。如果下面经常提到学校中数学教学的实际情况，那么，我的这些说明也并非基于对情况的不确定的印象，甚至基于我自己上学时的模糊记忆，因为我同席马克保持着经常的接触，他现在执教于哥廷根中学，经常向我报告数学教学的现状。确实，目前的数学教学已经比过去大大前进了一步。在今年冬天这一学期中，我要讨论"3A"(即 3 门主课：算术、代数和分析)，把几何放到明年夏天再讲。让我提醒一下，用中学的说法，这 3 门课一起叫做算术(指德国当时的情况——译者)，下面我要经常指出中学与大学中所用术语的不同，从这个小小的例子中你们也可以看到，只有保持生动的接触才能沟通相互的认识。

我要提到的第二本书是 H·韦伯和 J·韦尔斯坦因合编的三卷本《初等数学百科全书》[②]，在近年的出版物中，这本书最接近我的观点。在这个学期中，H·韦伯所撰第一卷《初等代数及分析百科全书》最为重要。不过我马上要指出这本书和我的讲授大纲的显著不

---

① F·克莱因：*Vorträge über den mathematischen Unterricht an höheren Schulen*，R·席马克整理。第一部分：*Von der Organisation des mathematischen Unterrichts*，莱比锡，1907 年以后提到时，这本书称为《克莱因—席马克合著本》。

② 第二版，莱比锡，1906 年。第四版出版于 1922 年，由 P·爱泼斯坦(P. Epstein)修订，以后提到时称为《韦伯—韦尔斯坦因本第一卷》。

中译者注：该书有郑太朴的中译本，共 3 卷(算术、代数、分析)，书名为《数学全书》，20 世纪 30 年代由商务印书馆出版。

同之处。在韦伯—韦尔斯坦因的书中，整个初等数学结构是用高年级学生成熟的语言系统地、逻辑地建立起来的，没有顾及学校里的实际出现次序。但是学校里的讲授应当顾及学生的心理，不应只讲究系统。可以说，老师必须是一个外交家，要考虑到小孩子的心理，以便抓住他们的兴趣；而且也只有用直觉上可以理解的方式讲授内容才能取得成功，到了高年级才可以用比较抽象的讲法作公理化的解释。举例来说，如果把数当作没有具体内容的抽象东西，并根据一些形式规则进行运算，那么小孩子就未必能理解。小孩子总是相反地把数的概念同具体的形象联系起来，这些具体的形象就是核桃、苹果之类好吃东西的数目，一开始只能，而且应当以那种看得见、摸得到的东西把数的概念向他们提出。这用不着多说，总之应当记住（当然不是处处如此），在各级教学中，甚至在大学中，都应当把数学同处在特定智力发展阶段上的学生真正感兴趣的东西联系起来，而且无论如何是要做到的。这正是近年来大学中努力注重应用数学的背景。中学里从来不像大学里那样忽视这种需要。我要在讲座中特别加以强调的，正是这种心理因素的价值。

韦伯—韦尔斯坦因和我的另一个不同之处是在确定中学数学教学的内容上。他们两位是倾向于保守的，而我持激进的态度。这些问题在我和席马克合著的书中已经详尽地进行了讨论。我们被大家称为改革派，希望把函数概念放到教学的中心地位，因为在过去两个世纪的一切数学概念中，凡用到数学思想的地方，函数概念总起着主导的作用。我们要尽快把这个概念引入教学，不断使用作图法，用 XY 系统来表示函数关系，而且事实上这种表示法今天在数学的一切实际应用中已被认为是自然而然的事。为了使这种改革成为可能，我们希望取消许多传统教学内容，即使这些内容本身可能很有趣，但从它们在现代文化中的地位来看，意义却并不重要。无论如何，强烈发展空间观念始终是首要考虑。不过达到了高级阶段就应该把教学远远推进到微积分初步上去，使自然科学家或保险专家在

上学时就掌握对他们不可或缺的工具。与这些比较新的观念相反，韦伯—韦尔斯坦因本质上坚守着传统的内容。在下面的讲座中，我当然要鼓吹新的观念。

我要提到的第三本书，是马克斯·西蒙(Max Simon)的《算术和数学教学法》这本很有启发的书[1]。他和韦伯及韦尔斯坦因一样，目前在斯特拉斯堡工作。西蒙的观点常常和我们保持一致，但他有时采取相反的立场。他是一个很主观，但又很热情的人，所以这些对立的观点叫他说起来就十分生动。举一个例子，自然科学家协会教育委员会建议在中学二年级有 1 小时几何预习，而目前一般在三年级开始。哪一个教学计划好，早就是有争议的问题，而学校里的习惯作法是经常变化的。教育委员会采取的立场即使再坏也不过是有待争论而已，但西蒙声称，这"比犯罪行为更坏"，又不提出一点点事实根据。他的书中这类段落很多。先于我这本书出版的书，有西蒙的《初等算术与代数分析相结合的教学法》[2]。

讲了这个简短的前言之后，我们就进入正题。前面已经提到，我将分 3 个大题来进行讨论。

---

① *Didaktik und Methodik des Rechuens und der Mathematik*，第二版，慕尼黑，1908 年。鲍梅斯特(Baumeister)的 *Handbuch der Erziehungs und Unterrichtslehre für höhere Schulen* 1895 年第一版的单印本。

② *Methodik der elementaren Arithmetik in verbindung mit algebraischer Analysis*，莱比锡，1906 年。

第一部分

算　　术

第一章 自然数的运算

让我们从算术的基础即正整数的运算讲起。就像以后各章一样,我们先提出中学里是怎样处理这些内容的,再讲从高等数学观点看它们意味着什么。

§1.1　学校里数的概念的引入

我只限于做一些简单的提示,这将使你们回忆起自己是怎样学到数的概念的。我这样讲的目的,当然不是像中学讲习班那样,为了把你们领进教学之门,而仅仅是为了摆出我们据以进行评论的材料。

教小孩学会整数的性质,学会整数的运算,再使他们彻底掌握,这是一个很难的问题,要他们下几年的工夫,从小学一年级学到 10 岁或 11 岁。德国的教法也许用直观和生成两个词来表达最为确切。也就是说,整个数的概念结构是在熟悉的、具体的事物的基础上逐步建立起来的,这与大学里学习用的逻辑及系统方法恰成鲜明的对照。

这一部分教学内容可以大致划分如下:小学一年级整整一年都学整数 1 到 20,前半学年从 1 学到 10。整数最初出现是用一个个点或一排排小孩熟悉的各种东西标上数字,然后用直观法讲授加法和乘法,使小孩牢记在心。

第二阶段教整数 1 到 100,引入阿拉伯数字,同时引入位制概念和十进制。附带说说,"阿拉伯数字"这个名称就像许许多多科学名称一样,是一个张冠李戴的名称。发明这种记数的形式实际上是印度人,而不是阿拉伯人。第二阶段的另一个主要目的是学会乘法表,

可以说必须要睡着了也背得出 5×7 或 3×8。当然学生要熟记乘法表到这种程度，这只有通过直观的手段，运用具体的东西使学生搞清楚之后，才能够说有把握。为此目的，常用算盘来帮忙。你们知道，它有 10 条柱子，一条在另一条之上，每一条上穿 10 个活动的珠子。适当移动这些珠子，就可以读出乘法的结果以及它的十进制记法。

最后在第三个阶段，教一位数以上的运算。运算都是一些已知的简单规则，其普遍正确性对于学生是不言而喻的，或者说是理所当然的。不过虽然说是理所当然的，但他们也不一定能把这些规则完全变成自己的规则，常常需要用权威的口吻灌输给他们，告诉他们如此这般，记不住不行。

这里我想再强调一点。这一点往往被大学教学所忽略，那就是联系实际生活去着重指出数的应用。从一开始学生就同取自实践的数打交道，离不开硬币、尺寸和重量。"这多少钱?"之类的问题，在日常生活中是极其重要的，一定要成为许多教学内容的中心。这样去学，很快就会达到解应用题的阶段，即必须深入思考才能确定用哪一种运算方法的阶段。再进一步，就要接触比例问题、混合运算问题等等。因此，除了上面我们用来概括这种教学性质的两个词——直观和生成以外，还可以加上第三个词即应用。

概括来说，数的教学目的也许可以归纳为：要使学生应用运算规则可靠无误，应以有关智力的平行发展为基础，而不必特别考虑逻辑关系。

这里我附带要你们注意：受过大学教育的教师和上过初等师范学校的教师的区别，由于所受教育的不同，往往造成学校教学上的不连贯。到了小学六年级或六年级以后，算术教师就要换了，上过大学的教师就要接替上过初等师范的教师，结果教学上的不连贯性往往会不幸地表现出来。可怜的小家伙们突然要去熟悉新的表达式，而不许用旧的了。最简单的一个例子是乘号的不同，小学教师要用"×"号，而上过大学的教师要用"·"号。诸如此类的矛盾当然是可

以克服的，只要教育程度较高的教师更多地关照自己的同事，在共同的基础上求得一致。如果你明白对于小学教师的工作必须予以多大的尊重，你就会觉得这是不难做到的。不妨设想一下，要把算术原则一遍又一遍地灌输给千千万万个没有学过的笨孩子，这需要多高级的教学法！用你在大学里学的那一套去试一下，十之八九会碰壁！

闲话少说，我们再回过来谈教学内容。我们要指出，过了中学三年级，特别是到了四年级[①]，算术就开始套上数学的高贵的外衣，转而采用字母符号来进行运算，这是转变期的一个特征。我们用 a, b, c 或 x, y, z 来表示任何一个数，起初只是正整数，并将算术的运算法则用到了字母所表示的数上，而这些数则没有具体的直观的内容。这是抽象化过程中的一大步。这一步就使我们有理由说：真正的数学是从字母符号的运算开始的。当然，这种转变不能突如其来，必须使学生自己逐渐习惯于这样明显的抽象。

似乎毫无疑问，为了搞好这一部分的教学，教师必须彻底了解运算的逻辑法则及基础，以及整数的理论。

§1.2　运算的基本定律

在考查支配加法和乘法的运算基本法则究竟是什么以前，人们早就熟悉了这些运算。运算的基本性质是 19 世纪 20 年代和 30 年代概括出来的，特别是英法数学家对此作了概括。不过这里我不准备细谈历史，如果你们想研究，我要像往常一样建议你们参考有关的百科全书，即伟大的德文版《数学及其应用百科全书》以及带有部分增订性质的法文译本[②]。无论哪座中学图书馆，即使只备一本数学

---

① 德国学制：四年小学、九年中学。因此，这里说的中学三、四年级相当于我国初一、初二。——中译者

② 由莱比锡杜伯纳出版社自 1908 年起出版到现在。第 1 卷已出齐，第 2 卷、第 3 卷接近于完成。法文版自 1904 年起由巴黎高特—维勒斯(Gauthur-Villars)出版社和莱比锡杜伯纳出版社合作出版，不幸自主编 J · 莫尔克(J. Molk)1914 年逝世后中断。

著作，也应该是这部百科全书，因为数学教师可以通过它进行所感兴趣的任何方向的研究。这里我们感兴趣的文章是第一卷[①]中的第一篇文章，H·舒伯特（H. Schubert）撰写的《算术基础》，由裘里斯·特纳里（Jules Tannery）和裘里斯·莫尔克（Jules Molk）译成法文。

回过来谈我们的正题。加法所依据的5个基本定律，我想列举如下：

(1) $a+b$ 仍然为一个数，即正数加正数总是可能的（减法却不同，在正整数的范围内不一定可能）。

(2) $a+b$ 是单值的。

(3) 结合律：$(a+b)+c=a+(b+c)$。

因此完全可以脱去括号。

(4) 交换律：$a+b=b+a$。

(5) 单调律：若 $b>c$，则 $a+b>a+c$。

关于这些性质，只要回忆一下计算的过程就立刻可以弄明白，但是必须正式提出，以便能从逻辑上看出以后发展的合理性。

乘法有5个与加法完全相似的规律：

(1) $a\cdot b$ 仍然为一个数。

(2) $a\cdot b$ 是单值的。

(3) 结合律：$a\cdot(b\cdot c)=(a\cdot b)\cdot c=a\cdot b\cdot c$。

(4) 交换律：$a\cdot b=b\cdot a$。

(5) 单调性定律：若 $b>c$，则 $a\cdot b>b\cdot c$。

乘法和加法混合运算还服从以下法则：

(6) 分配律：$a\cdot(b+c)=a\cdot b+a\cdot c$。

很容易看出，一切初等运算都可以依据这11个法则。这一点，只要用一个简单的例子就足以说明。就拿 7×12 来说，根据分配律得：

---

① 该卷名为《算术与代数》，W·Fr·梅耶（W. Fr. Meyer）(1896—1904年)主编，法文版主编为J·莫尔克。

$$7 \times 12 = 7 \cdot (10 + 2) = 70 + 14。$$

若将 14 分成 10 + 4（即“逢十移出”），则根据加法结合律，得

$$70 + (10 + 4) = (70 + 10) + 4 = 80 + 4 = 84。$$

从这个运算过程，可以认清一般十进制运算的步骤。你们最好去想出一些比较复杂的例子。可以概括地说：一般的整数运算就是结合所记住的加法表和乘法表，反复运用上述 11 个基本法则。

那么，哪里用到单调性定律呢？在一般的形式运算中，说实在的，该定律是多余的，但在解某些问题时并非如此。我提醒你们注意十位数的快速乘除运算[①]。这种运算具有极大的实用价值，不过遗憾的是中学生知道它的太少，大学生当中知道的也不多，尽管德国中学二年级有时提到。举例来说，假定我们要计算 567 · 134，再假定此两数的个位数字的精确性有问题，譬如说是物理测量的结果。这样就没有必要求精确的乘积，因为不能保证有精确的结果。不过必须知道乘积的数量级，即精确值在十位数或百位数。依据单调性定律就可以立刻得出这种估计，因为据此定律，所求值在 560 · 134 和 570 · 134 之间，或在 560 · 130 和 570 · 140 之间。我把细节留给你们自己去考虑，你们至少会了解，单调性定律在快速运算中是经常用到的。

系统地讲解这些基本法则当然不在中学考虑之列，但在学生对数的运算有了具体的了解并已掌握牢固之后，准备过渡到字母符号运算的时候，老师就应该借机会叙述一下，至少叙述一下结合律、交换律及分配律，并举出许多明显的数字例子来加以说明。

§1.3　整数运算的逻辑基础

中学的数学教学当然不会提到更难的问题，不过当前的数学研

---

① 单调性定律以后在无理数理论中还会用到。

究实在是从下述问题开始的，这些问题就是：我们怎样去论证上述的基本法则？究竟怎样去解释数的概念？根据我在讲这些课之前已经声明的目的，我要把这个事情解释一下，以便竭力从另一个观点去观察，使我们对中学课程有一个新的认识。我们之所以愿意这样做，是因为在大学几年中现代的数学思想从四面八方挤到你们的脑子里来，但并不一定同时对这些思想的心理学意义做任何说明。

首先，就数的概念而言，它的起源是很难弄清楚的。这些最难懂的东西，不去管它们也许是最快乐不过了。对于哲学家们讨论得如此认真的这些问题，如果要得到比较全面的资料，我必须建议你们去读前面已经提到过的、收在法国《百科全书》里的那篇文章，这里我只提上几句。一个普遍接受的信念是，数的概念同时间概念、时间交替概念有密切的关系。哲学家康德(I. Kant)和数学家哈密顿(W. R. Hamilton)是这种观点的代表。其他一些人则认为，数的概念同空间概念更有关系。这一派人把相邻几个物体同时进入知觉的心理过程当作数的概念的基础。但还有一些人从数的概念中却看到了独立存在于空间和时间并与之协调，甚至超越其上的某种独有的心智官能。我认为这种观点不妨引用《浮士德》里的两行诗来概括，闵可夫斯基(H. Minkowski)在其所著的《丢番图近似》一书的序言中也曾用这两行诗来说明数：

"Göttinnen thronen hehr in Einsamkeit,

Um sie kein Ort, noch weniger eine Zeit."

(巍巍御座，女神独在

周无空间，遑论时间!)

固然，这个问题主要涉及心理学和认识论问题，但我们所说的11个法则的论证，是离不开逻辑学上种种问题的。至少近年来对于这些法则的相容性的研究表明是如此。下面我们来分析4种观点。

(1) 根据以康德为代表的第一种观点，运算法则是知觉的直接而必然的结果，而"知觉"这个词应从最广义的角度来理解为"内知

觉”或直觉。据此，数学不应理解为处处建立在可以用实验控制的外界经验事实基础上的科学。举一个简单的例子，交换律的建立，就是因为观察到了这样一个相关的图形：$\begin{smallmatrix}\bullet&\bullet&\bullet\\\bullet&\bullet&\bullet\end{smallmatrix}$。图上有两排点，每排各有 3 个点，这也就是：$2 \cdot 3 = 3 \cdot 2$。可能有人会出来反对说，点数只要适当多一点，这种直接知觉就不行了。答复是：我们可以求助于数学归纳法，即：若一假设对于小正整数成立，并设其对正整数 n 成立，总能证明其对于 $n+1$ 亦成立，则此假设对每一正整数皆成立。这个法则我认为真是一个直觉的真理，它使我们超越感官知觉所达不到的界限。这个立场大致就是庞加莱(H. Poincavé)那些著名哲学文章中的立场。

如果我们理解这个问题对于 11 个基本运算法则何以成立的根据至关重要，那么请记住：包括算术在内，整个数学都是建立在这 11 个法则的基础上的。因此，根据刚刚加以概括的、对于运算法则的观念，可以说整个数学结构的可靠性都是建立在直觉的基础上的，这并不过分武断。不过“直觉”这个词应该从最广义的角度来理解。

(2) 第二个观点是第一个观点的另一种说法。据此观点，我们可以竭力把 11 个基本法则分成大量较小的步骤，只要直觉地从其中直接取一个最简单的步骤，其余步骤就可以按照逻辑法则推导，不必再用直觉。在前面，逻辑运算是在 11 个基本法则确立之后才开始的，现在则可以早点开始，即在选出比较简单的法则之后就可以。这里，直觉和逻辑的界限的位置变动了，而且让逻辑占了上风。1861 年，赫尔曼·格拉斯曼(H. Grassman)写了一本《算术读本》，朝这个方向作了开创性的工作。举其一例，我只提一下，交换律可以借助于数学归纳法原理从结合律推导出来。由于叙述精确，你们可以在格拉斯曼这本书的旁边放一本意大利人皮亚诺(G. Peano)写的 *Arithmetices Principia Nova Methodo Exposita*。不过别看了这个标题以为这本书是用拉丁文写的！它是用作者独创的符号语言写的，创立这种符号的目的是表示各个逻辑步骤，用以强调仅此而非

他。皮亚诺想以这种方式来保证他仅仅使用了他特别提出的原理，而把直觉的因素完全排除掉。他想，如果他使用日常语言，就会有无数不可控制的观念联想及知觉的暗示不知不觉地掺杂进来，而他想避免的正是这种危险。还要注意，皮亚诺是意大利一大学派的领袖，这一派人就是企图以类似的方式把各个数学分支的前提分成一个个小的步骤，并借助于那种符号语言去研究种种前提的真正逻辑关联。

(3) 现在我们来谈这些数学思想在现代的发展，不过这种发展也是受到皮亚诺的影响的。我这是指把点集理论放到突出位置来解释算术基础的方式。如果我告诉你，不但线段上的所有的点，而且所有的整数都是点集的特例，你就会对点集理论涉及面之广有一个认识。正如普遍了解的那样，康托(G. Cantor)是第一个把这个普遍观念变成有条理的数学思维对象的人。对他所创立的点集理论，现在年轻一代数学家是非要深刻注意不可的，以后我们还要使你们对它有一个粗略的看法。至于现在，只要把根据它建立的、新的算术基础说一下，指出其发展趋向就够了。这一点可以概括如下：整数及整数运算的性质要从点集的一般性质及抽象关系来推出，以便使算术的基础尽可能完整可靠并带有普遍意义。走这条路的开创人物之一是理查德·戴德金(R. Dedekind)，他在一本虽小但极为重要的书中曾经作过尝试，想要为整数建立这样一个基础。他那本书名叫 *Was sind und was sollen die Zahlen?*[①]。H·韦伯在他和韦尔斯坦因合著的《初等代数与分析》第二版第一卷第一部分中倾向于这个观点。不过推论相当抽象，仍存在着一些严重的难点，所以韦伯在第三卷附录[②]中只用有限点集作了比较初等的叙述。在以后几版中，这个附录被收入第一卷。你们当中对这种问题感兴趣的人，特请查阅那个叙述。

(4) 最后，我要提到数的纯形式的理论，这个理论确实应追溯到

---

① 布朗施魏格(Braunschweig)出版社，1888 年，第三版，1911 年。

② *Angewandte Elementarmathematik*，H·韦伯修订，莱比锡，1907 年。

莱布尼兹(G. Leibniz)，而后来希尔伯特又把它提升到突出的地位。1904 年他在海德堡数学大会上的演讲“关于逻辑和算术的基础”①，对于算术是很重要的。他的基本观点大致如下：只要有了 11 个基本运算法则，就可以用字母符号 a, b, c, …来进行运算，这些符号实际上代表着任意整数，但你心里不必记着它们有实际的数的意义。换句话说，可以设 a, b, c,…为没有具体意义的东西，或对其意义一无所知。只要我们同意，可以根据那 11 个法则把它们加以组合，但这些组合又不一定要具有任何已知的实际意义。显然，这样就可以完全像平常用实际数字进行运算一样，用 a, b, c, …来运算。这里只有一个问题：这样运算会不会导致矛盾？好吧，平常我们说直觉告诉我们是有数存在的，而且 11 个法则对这些数成立，所以就不可能有潜在的矛盾。不过，既然现在我们不认为符号有确定的意义，那就不允许诉诸直觉。所以事实上这里的问题是一个全新的问题，是要我们从逻辑上证明，对我们的符号，建立在 11 个基本法则基础上的任何运算绝不会导致矛盾。也就是说，这 11 个法则是一致的或相容的。我们在讨论第一派观点时采取的立场是：数学的可靠性在于存在着与其定理相适应的直觉的东西，但拥护形式立场的一派却必须这么讲：数学的可靠性在于它能证明：从形式上考虑的，并不顾到直觉内容的基本法则是一个逻辑上一致的系统。

下面我发表几点意见，作为这一部分讨论的结语：

(a) 希尔伯特在海德堡演说中提了这一切观点，但他对任何一个观点都没有完全追随到底。后来他在一本教程中把那些观点推到更远的地步，但随后又放弃了。我们可以这样说，这是一个可以研究的领域。

(b) 在我看来，把直觉完全挤出去，以求取纯而又纯的逻辑研究，并不完全可行。我觉得，最低限度也要保留一点直觉。即使用符

---

① “Verhandlungen des 3. internationalen Mathematiker - kongresses in Heidelberg”, 8 月8—13 日，1904 年，第 174 页及以后数页，莱比锡，1905 年发表。

号运算，也总是要在最抽象的公式处理中利用一定的直觉，哪怕只想到字母符号的形状，以便再认出符号。

(c) 退一步说，假定所提出的问题已经没有争论地解决了，11 个基本法则的相容性也已从逻辑上证明了，即使如此，也还有讨论的余地，我想加以指出，并予以最大的强调。我们必须清楚地看到，以纯形式方式建立起来的算术、整数理论，是既不曾有过，以后也不可能有的。不可能以纯逻辑方式证明，以那样方式建立起来的相容的法则，对于我们直觉上所熟悉的数量是真正成立的。此外，也不可能证明我们所说的未定义的，并对之进行运算的东西，就是真实的数，就是直觉上意义很清楚的加法和乘法。已经取得的成就，毋宁说是在于：它把建立算术基础的这个复杂的攻不下的问题分成了两个部分，而且使第一部分——纯逻辑问题，即建立独立的基本法则或公理及研究其独立性和一致性问题已经有了探讨的可能。第二部分更多地属于认识论问题，涉及这些法则应用于现实情况的论证，至今甚至还没有触及。当然如果真的要建立算术的基础，也是非解决不可的。这第二部分问题，本身是极其深刻的，其困难在于一般认识论领域。我可以用多少近乎悖论的方式清楚不过地说明：任何人若在纯数学研究中仅仅容忍纯逻辑的话，为了自圆其说，都不得不考虑算术基础问题的第二部分，也就是把算术本身看做是应用数学的领地。

由于在这一点上经常产生误解，因此我感到有必要在这里很详细地讨论一下，因为人们简直忽视第二部分问题的存在。这绝不是因为我同希尔伯特本人的争论，如果基于这样的假设，那就既不能正确理解我同他的意见分歧，也不能正确理解我同他的意见一致之处。

耶拿市的托梅造过一个言简意赅的词语——“无思想的思想家”，用来指仅仅对空无意义的事物进行抽象研究的人。这种人只限于研究不说明任何问题的空理，不仅忘了上述第二部分问题，而且往往也忘了数学中的其余的一切。这个开玩笑似的说法，当然不能用来指进行抽象研究，同时也研究其他许多不同性质问题的人。

结合以上的简评,我提出几个一般的问题,以引起你们的注意。许多人认为教一切数学内容都可以或必须从头到尾采用推导方法,从有限的公理出发,借助于逻辑推导一切。某些人想依靠欧几里得(Euclid)的权威来竭力维护这个方法,但它当然不符合数学的历史发展情况。实际上,数学的发展是像树一样的,它并不是有了细细的小根就一直往上长,倒是一方面根越扎越深,同时以相同的速度使枝叶向上生发。撇开比喻不说,数学也正是这样,它从对应于人类正常思维水平的某一点开始发展,根据科学本身的要求及当时普遍的兴趣的要求,有时朝着新知识方向进展,有时又通过对基本原则的研究朝着另一方向进展。例如,我们今天对于数学基础的立场,不同于几十年以前;我们今天可能当作最终原则来叙述的东西,过了一段时间也必然会被超越,因为今天看来最新的真理会得到更为过细的分析,又需要归结为更一般的东西。由此我们可以明白,对于数学中的基础研究来说,是不存在最终的终点的,也不存在最初的起点,来为数学教学提供绝对的基础。

再谈一点关于数学的逻辑和直觉之间的关系、纯数学和应用数学之间关系的意见。我已经强调过,在小学里算术的教学从一开始就伴随着应用,小学生学习运算规则不仅是为了理解它们,而且是为了用它们解决什么问题。数学教学就应该永远是这样的!当然,逻辑关系,或可以说数学机体上的硬骨架,必须保持下去,以便使数学具有它所特有的可信性。但是数学的生命,数学的最重要的动力,数学在各方面的作用,却完全有赖于应用,即取决于那些纯逻辑内容和其他一切领域之间的相互关系。把应用拒之于数学门外,就等于只从骨架中去找活生生动物的活力,而不考虑肌肉、神经和组织,不考虑动物的本能,总之就是不考虑动物的生命本身。

在科学研究中,确实常有纯科学和应用科学的分工,但即使如此,如果想使情况良好,就仍然要另外做一些约定使之保持联系。不过无论如何应该特别强调,在学校里要求这样的分工,要求一个教师

这样地专门化是不可能的。往极端里说,如果某学校指派一位教师把数教成没有意义的符号,指派第二位教师去在符号和实在数字之间搭上桥,指派第三、第四、第五位教师去讲这些数字在几何学、力学、物理学上的应用,这些教师都压在学生头上,这样的组织教学是不可能的。这样的话,教学内容就不可能使学生理解,各个教师恐怕也甚至不能互相理解。而学校教学本身恰好需要各个教师成为某种程度上的多面手,对最广义的纯科学及应用科学领域有一个大致的了解,以便对科学分工太细的情况采取一个理想的补救办法。

为了针对前面的话提出一个实际办法,我再来引用一下我们在引论中提到过的德累斯顿建议书。我们在那篇建议书中直陈,由于应用数学从 1898 年后已经成为优秀教师考试中的一个专门项目,因此应该把它列为一切普通数学教学的必修课程,以便使学生能始终兼教纯数学和应用数学。此外应指出,教育委员会的密伦教学大纲[1]也已宣布下列 3 个任务,作为上学年(指 1923 年——译者)数学教学的目的:

(1) 科学地概述数学的系统结构。

(2) 掌握处理数值问题及作图问题的一定技巧。

(3) 认识数学思想对于自然科学及现代文化的重大意义。

对于这一切规定,我深表赞同。

---

① *Reformvorschläge für den mathematischen und naturwissenschftlichen Unterricht, überreicht der Versammlung der Naturforscher und Ärzte zu Meran*,莱比锡,1905 年。并可参阅 *Gesamtbericht der kommission* 中重印之文(第 93 页),以及我和席马克合著《数学教学组织》一书的第 208 页。

第二章
数的概念的第一个扩张

讲完前一章数的运算，这一章就来讨论数的概念的扩张。在中学里，通常依次采取下列程序：

(1) 分数和分数运算的介绍。

(2) 结合字母符号运算，开始学习负数。

(3) 通过不同的实例比较完整地提出无理数的概念，然后逐步地导入实数连续统的概念。

前两点按什么程序讲是无所谓的，这里先讲负数，再讲分数。

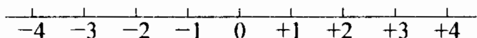
§2.1 负　　数

我们先来讲讲术语上的问题。在中学里，正数和负数统称为相对数，以区别于绝对数（正数），但是在大学里这种说法不普遍。还有，中学里“代数数”同相对数互称，而“代数数”这个术语在大学里意义就很不同了。

建立负数概念的原因，是因为要求在一切情况下都有可能进行减法运算。如果 $a<b$，那么 $a-b$ 在自然数范围内就是没有意义的，但数 $c=b-a$ 确实存在，所以写

$$a-b=-c,$$

称之为一个负数。根据这个定义，我们立刻可以在横坐标轴上作一些等距离的点，由原点向两个方向扩展，通过等距离点的刻度来表示一切整数。



今天所有受过教育的人都可以看懂这个图，也许可以假定，主要由于温度计而得到这种普遍的了解。商业资产负债表上的借方和贷方，同样可以对负数提供一个熟悉的图表形式。

但是我要立刻着重指出，中学里负数概念的引入，在原则上是极为困难的一步。学生已习惯于直观形式，即通过事物的具体数量来表示数，现在他们会觉得运算的符号和结果与以前太不同了。负数对他们是某种新的概念，与他们从具体事物的数量得来的观念没有共同之点，但是他们不得不把负数当作实有其物来运算，尽管要比过去抽象得多。这是由具体数学向形式数学的第一次转折。要完全掌握这种转折中出现的问题，需要有高度的抽象能力。

下面来详细探讨引入负数后所发生的运算问题。引人注目的第一件事是加减运算打通了，即：加一个正数就是减去它的相反数。马克斯·西蒙有一个滑稽的说法，他说创立负数是为了有可能毫无例外地进行减法运算，但由于负数的创立，减法运算就不再独立存在了。

对于正负数范围内的这种新的加法运算（包括减法运算），前述5个形式法则仍然成立。这5个法则简单来说就是：

（1）运算恒可进行。

（2）单值。

（3）结合律。

（4）交换律。

（5）单调性定律。

注意：结合法则5，$a<b$ 现在用几何表示就是 a 在左、b 在右，因此举例来说，$-2<-1$，$-3<-2$。

正负数相乘中的注意点是符号规则，即

$$a \cdot (-c) = (-c) \cdot a = -(a \cdot c),$$

以及

$$(-c)(-c') = +(c \cdot c')。$$

特别是后一个规则——“负负得正”，往往是危险的绊脚石。下面马上就回过来讲这些规则的内在含义，现在就用一句话来定义一串正数和负数的乘法：乘积的绝对值等于诸因子绝对值的乘积，乘积的符号是正或是负，随负因子个数为偶数或为奇数而定。据此，正负数范围内的乘法又具有下列特殊的性质：

(1) 运算恒可进行。

(2) 单值。

(3) 结合律。

(4) 交换律。

(5) 对于加法满足结合律。

只有单调性法则有变化，代之以下一规律：

(6) 若 $a > b$，则随 $c \lessgtr 0$，有 $a \cdot c \lessgtr b \cdot c$。

现在我们要问：再次从纯形式考虑的这些法则是否相容？必须立刻承认，其纯逻辑证明的可能性比整数情况要小得多。只可能作一个约定，即若上述法则对于整数成立，则对于负数也成立。但在整数得到逻辑相容性证明以前，不得不承认上述法则的相容性纯系基于这样一个事实，即有一些直觉的东西，它们有一种服从这些法则的直觉关系。上面我们已举出横坐标上整数点序列的例子，现在只要说明那种运算规则的含义：加式 $X' = X + a$，其中 a 是确定的，对应于每一点 X 有第二点 X'。这样，无限直线只要根据 a 的正负向左或右移动量 a。类似于此，乘式 $X' = a \cdot X$ 也是直线的相似变换，对于 $a > 0$ 为单纯的伸缩，对于 $a < 0$ 为伸缩与关于原点之对称变换的复合。

现在允许我来讲讲这一切的历史发展过程。你们不要以为负数是某一个聪明人的发明，以为是他从几何表示形式中悟出了负数的

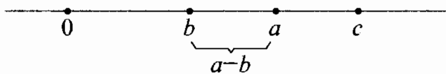
道理，或许也就同时确立了负数的逻辑相容性。不是的！负数是经过很长一段发展时期才夺路而出，使数学家不得不用到它。只是到了 19 世纪，在人们已经用它进行运算好多个世纪以后，才考虑到了它的逻辑相容性。

讲起负数的历史，我先要提到：不像许多人所想的那样，样样事情希腊领先。古希腊人肯定没有负数。这个发明应归功于印度人，印度人还创造了我们用的十进制，特别是数零。欧洲人是在文艺复兴时代逐渐用起负数来的，当时刚刚过渡到用文字符号运算的阶段。这里也不能不提，文字符号的运算是维塔（Vieta）在他写的 *In Artem Analyticam Isagoge*（1591 年）一书中所创立的。

从今天的观点来看，我们已有所谓正数运算的括弧规则，假如把相应的减法法则也包括在我们的基本公式之内的话，那么括弧规则当然也包括在基本公式之内。不过我想通过两个例子把括弧规则谈得再详细一点，以表明它们有极其简单的直觉证明的可能。按照印度人的习惯，只需把图画出来，说一声“请看！”就够了。

(1) 给出 $a>b$ 及 $c>a$，其中 a, b, c 皆为正数。那么，$a-b$ 为一正数，小于 c，即 $c-(a-b)$ 必作为正数存在。现在用横坐标把数字表示出来，表明点 a 和点 b 之间的线段具有长度 $a-b$。看一下图示就可以明白，如果从 c 段中取走 $a-b$，结果同我们先取走整个线段 a，再放回 b 段一样，即

$$c-(a-b)=c-a+b。\tag{1}$$



(2) 给出 $a>b$ 和 $c>d$，$a-b$ 和 $c-d$ 为正整数。我们来看乘积 $(a-b)\cdot(c-d)$。为此作一以 $a-b$ 及 $c-d$ 为两边的矩形（标以斜线阴影）（见图 2.1），此矩形的面积即为所求的数：$(a-b)\cdot(c-d)$，并且此矩形是以 a 和 c 为边的矩形的一部分。为了从后一矩形

中求出前一矩形，先取走标以水平线阴影的矩形面积 $a \cdot d$，再取走标以竖直线阴影的矩形面积 $b \cdot c$。这样已两次取走双重线阴影的矩形面积 $b \cdot d$，必须把它放回。这正是已知公式

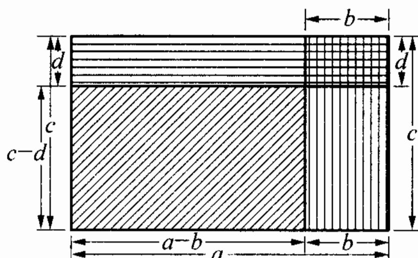


图 2.1

$$(a-b)(c-d)=ac-ad-bc+bd。\qquad (2)$$

在字母记号运算的基础上导入负数，其中所涉及的最重要的心理活动，是人类本性的一般表现，因为人类不由自主地倾向于在更一般的情况下运用一些法则，而不顾这些法则只是在一些特例下导出并成立的。这就是赫尔曼·亨克尔(Hermann Hankel)在《复数系理论》一书中以《形式法则的承袭原则》为题首先发表的算术指导原则。我愿向你们推荐这本极有趣的书。运用到负数发展的这个特定情况上来，上述原则等于是说，在(1)式和(2)式这类公式中，我们倾向于忘却关于 a 和 b 相对大小的给定的假设，而把它们运用于另一些情况。例如：如果把(2)式应用于 $a=c=0$（原公式在这种情况成立与否，压根儿未曾加以证明），便得到 $(-b)\cdot(-c)=bc$，即负数相乘的符号法则。这样，我们事实上几乎不自觉地不经假设即已导出一切规则，而循前思路，这些规则本来必须看做必要的假设，才能使旧的法则对于新的概念也成立。说实在的，过去的数学家对于这种抽象也不满意，他们有时称负数为虚构数、假数之类，就表明了他们于心不安。尽管有顾虑，但负数在16世纪、17世纪还是得到了越来越多的人的公认，因为它很有用。对于这一点，解析几何的发展无疑起到了很大的促进作用。话说回来，疑问还是存在着，而且只要人们继续想把它用事物的个数概念来表示，而没有认识到，在新概念建立之后，其逻辑形式法则起主导作用，这种疑问也就会存在下去。正由于

这个原因，人们曾反复地企图证明符号法则。19 世纪提出了一个简单解释：谈论定理的逻辑必要性是没有用的，换句话说，符号法则不能证明；人们只关心这个法则在逻辑上是否允许。同时 19 世纪还提出解释说，这些法则是任意的，取决于使用上的方便，例如受承袭性原则的制约。

在这一点上，我们不禁会经常想到，事物的发展有时比人更有理。想一想：数学上最伟大的进展之一——负数及负数运算的引入，竟不是某一个人自觉的逻辑思考的创造。相反，它的缓慢的、有机的发展，是与事物广泛地打交道的结果，所以几乎好像是字母记号的运算把负数教给了人。过了很长一段时间，人才有了理性的认识，知道已经发现了某一正确的、与严格的逻辑相容的法则。不过无论如何，对于新概念的建立来说，纯逻辑的功能仅起规定的作用，而从来不起唯一的指导原则的作用，因为总还有其他许多概念体系满足逻辑上相容性的要求。

如果你们想了解一下有关负数历史问题的其他文献，我向你们推荐特罗夫克(Tropfke)的《初等数学史》[①]。这是一本极好的资料集，以流畅的表述提供了大量有关初等数学概念、观点、术语发展的细节。

如果我们现在带着批判的眼光去看中学里负数的教法，常常可以发现一个错误，就是像老一代数学家如上指出的那样，努力去证明记号法则的逻辑必要性。他们从 $(a-b)\cdot(c-d)$ 的公式导出 $(-b)(-d)=+bd$，以为就得到了证明，完全忽略了这个公式之所以成立取决于不等式 $a>b,\ c>d$。因此，证明是虚假的，本来可以根据心理学的考虑通过承袭性原则而得出法则，现在却让位于一种伪逻辑的考虑。学生第一次听到这样逻辑证明时，当然是听不懂的，

---

① 两卷本，莱比锡，1902 年、1903 年。第二版已修订，大大增加了篇幅，变成 7 卷本，其中 6 卷已于 1924 年出版。还可以参考 F·凯乔里(F. Cajori)著，麦克米伦(Macmillan)公司出版的《数学史》。

而最终只好相信；如果在高年级再讲的时候，还不能使学生形成正确的概念，那么某些学生就会产生一种根深蒂固的观念，以为整个概念是神秘而不可理解的，但事情竟常常如此。

我反对这种做法，我请求你们一般地不要把不可能的证明讲得似乎成立。大家应该用简单的例子来使学生相信，或有可能的话，让他们自己弄清楚：从实际情况来看，承袭性原则所包含的这些约定关系，恰好是适当的，因为可以得到一致方便的算法，而其他任何一种约定，总要强迫我们考虑许多特例。确实必须不急不躁，让学生有时间在接受这种知识后思想转过弯子来。尽管很容易说明白，其他的约定都是不好的，但必须向学生强调，普遍有用的约定确实存在，这个事实真是奇妙之至！同时使他们明白，这绝不是不言而喻的。

关于负数理论的讨论，我就说到这里为止。下面请大家同样来考虑数的概念的第二个扩张。

§2.2　分　　数

让我们从中学里的分数的处理谈起。在中学里，分式$\frac{a}{b}$从一开始就有完全具体的意义。与整数的图式表示相对照，只有依据的不同：从事物的数目转到了事物的量度，从可数事物的讨论转到了可量度事物。例如，币制或重量系统（附加上某些限制条件）是可量度的例子，而长度系统则是地道的可量度的系统。这些例子都可以给每一个学生以分数的概念。没有什么人对于领会$\frac{1}{3}$米和$\frac{1}{2}$磅的意义会有多大的困难。分数之间＝，＞，＜的关系，是可以通过同样的具体直觉马上建立起来的。同样，分数的加减运算，以及分数乘以整数的运算，也是如此。在这之后，一般的乘法可以很容易得到理解：用$\frac{a}{b}$乘一数，就是乘以 a 再除以 b。换句话说，乘积之由被乘数导出，

正如$\frac{a}{b}$之由 1 导出一样。然后分数的除法就教成乘法的逆运算：a 除以$\frac{2}{3}$等于这样一个数，此数乘以$\frac{2}{3}$等于 a。这些分数运算概念与负数概念合起来，就最终得出了一切有理数的全体。我不能详细说明这个形成过程，在中学里当然要用相当长的时间。我们最好立刻把中学里的教法同现代数学完美的表述作一比较，为此可用韦伯—韦尔斯坦因和伯克哈特(Burkhardt)的书作参考。

韦伯—韦尔斯坦因主要强调形式观点，他从种种可能的解释中选出一种对各种解释都是必然相通的解释。根据这个观点，分数$\frac{a}{b}$是一个记号，是一个可以根据某些法则进行运算的一个“数偶”。这些法则在我们上面的讨论中是自然而然地从分数的意义上得出的，但这里却有任意的约定的性质。例如，分数的分子和分母同乘以或同除以相同数的定理，学生视之为再明显不过的定理，在这里却作为相等的定义：当 $ad=bc$ 时，分数$\frac{a}{b}$，$\frac{c}{d}$称为相等。大于及小于亦可类似地给出定义，并且将两个分数$\frac{a}{b}$，$\frac{c}{d}$之和定义为$\frac{ad+bc}{bd}$，等等。从而证明：按上述方式定义在新的数域中的运算，在形式上正好具备整数乘除的性质，即满足一再举出的 11 个基本法则。

伯克哈特没有像我们已简要介绍的韦伯—韦尔斯坦因的表述那样强调形式观点。伯氏把分数$\frac{a}{b}$看做整数域内的两个运算的序列——乘以 a 及除以 b，运算的对象是任意选择的整数。如我们采取$\frac{a}{b}$，$\frac{c}{d}$这样两个“运算偶”，那就对应于分数的乘法。很容易看出，这样的运算无异于乘以 $a\cdot c$ 并除以 $b\cdot d$。所以分数乘法法则

$$\left(\frac{a}{b}\right)\cdot\left(\frac{c}{d}\right)=\left(\frac{a\cdot c}{b\cdot d}\right)$$

是从明明白白的分数的意义中得出的,但不仅仅是任意的约定所确定的。我们当然可以同样对待除法。另一方面,加法和减法就不能借助于分数的这种表示法作那样简单的解释,所以公式

$$\frac{a}{b}+\frac{c}{d}=\frac{(ad+bc)}{bd}$$

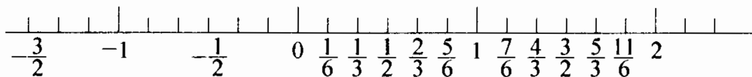
对于伯氏也仍然是一个约定,对于这个约定他只提出似然成立这样的理由。

现在我们来把中学里的旧的表述同刚才简要介绍的现代的概念作一比较。根据现代的概念,无论是哪一本书,不说数的概念的扩张,讲的内容确实不出整数的范围,仅仅假设整数的全部已直觉地掌握,或已知其运算法则,新的对象定义为数偶或整数的一些运算,这些都完全套在整数的框架以内。另一方面,中学里的处理方法是完全建立在新获得的可量度量的概念上的,并借这个概念对分式提供一个直接的直觉形象。如果我们假设有一个人具有整数概念,但无可量度的量的概念,那么就能更好地理解这种区别。那个人对中学里的说法可能完全不能理解,但却能很好地理解韦氏和布氏的讨论内容。

两个方法哪一个好呢?各有什么利弊?答复是:正像对不同的整数概念提出的类似的问题一样,现代的表述当然完美一些,但是欠丰富,因为传统课程中当作一个完整对象来讲的东西,它实际上只讲了其中一部分:即对原来称"分数"的某些算术概念及"分数"运算作抽象的、逻辑上完备的介绍。但它对一个完全独立的、同样重要的问题没有作出解释:我们能不能真的把这样导出的理论学说运用于具体的可量度的量?我们可以再称这个问题是"应用数学问题",并可作完全独立的处理。说实在的,这样分割开来,在教育学上是不是好,是成问题的。韦氏更是把这个问题典型地一劈两半。截此为止我们所介绍的只是他们对分数的抽象的引入。然后,韦氏和伯氏用专门一部分(第五部分:比率)去论述有

理数与外在世界的关系问题。他们的论述当然十分抽象，超过直觉所能接受的程度。

现在对有理数的整体谈点一般意见，以此来结束分数的讨论。为了清楚起见，我要利用直线表示形式。设想所有具有有理坐标的点都已在这条直线上一一标出，并简称这些点为有理点。于是我们说坐标轴上这些有理数点是“稠密的”，意思是每一个区间不管多小都有无限多的有理数点。如果想避免把任何新的东西放进有理数的概念里去，那么可以用更抽象的方式说，在任何两个有理数点之间总是存在另一个有理数点。由此可以从有理数点的整体中分出一些有限的部分，它们既没有最小的也没有最大的元素。在 0 和 1(包括 0 和 1)之间的有理数点的全体，是一个例子，因为在 0 和 1 之间给出任何一数，在这个数和 0 之间总会有一个数，即较小的数，在这个数和 1 之间也总会有一个数，即较大的数。这些概念系统发展的结果，就是康托的点集理论。实际上，我们以后将用到有理数的全体，连同刚才提到的性质，作为点集的一个重要例子。



下面来谈数的概念的第三个扩张：无理数。

§2.3 无 理 数

这里我们就不花时间去讨论中学里是怎么处理的了，因为在中学里顶多是讲几个例子。让我们马上来谈历史发展。从历史上来看，无理数概念的起源当然在于几何直觉，在于几何学的要求。如果像刚才认为的那样，认为有理数的集合在坐标轴上是稠密的，那么在坐标轴上还存在着另一些点。毕达哥拉斯(Pythagoras)据说曾用大致如下的方式表明这一点。给出一直角三角形，勾股长度为 1，则斜

边的边长为$\sqrt{2}$(图 2.2),这当然不是一个有理数;因为如果置$\sqrt{2}=\frac{a}{b}$,其中a和b是互素的整数,那么根据整数的可除性法则就很容易导致矛盾。如果我们现在在坐标轴上从0开始截取斜边长的那一段,就得到一个非有理数点,那个点不是在坐标轴上稠密的原来集合中的一个点。进一步,毕达哥拉斯肯定知道,在绝大多数情况下,边长为m,n的直角三角形,其斜边长$\sqrt{m^2+n^2}$是无理数。据说为了庆祝发现这个极为重要的事实,毕达哥拉斯拿出100头牛来献祭,这确实是值得的。我们也知道,毕达哥拉斯学派喜欢寻求使得直角三角形三边成为可公度的那些特殊的数对m及n,如果选择适当的测度单位,则这时直角三角形3边之长皆可以用整数表示(所谓毕达哥拉斯数)。毕达哥拉斯数的一个最简单的例子是3,4,5。

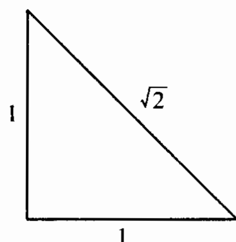


图 2.2

后来,除了这些最简单的无理数外,古希腊的数学家们还研究了比较复杂的无理数,如欧几里得式的$\sqrt{\sqrt{a}+\sqrt{b}}$,等等。但一般地说,他们主要限于反复开方求得的,因而可以用直尺和圆规作出几何图形来的无理数。无理数的一般概念,他们还不知道。

不过我必须多少修正一下这个说法以免产生误解。更精确的说法是:古希腊人没有掌握根据有理数以算术方式提出或定义一般实数的方法。这个方法是现代数学发展的结果,下面我们很快就要介绍。不过从另一个观点来看,他们是熟悉一般实数(不一定是有理数)的概念的,但是这个概念对他们来说表现形式完全不同,因为他们还不会用字母符号来表示一般的数。他们确实研究过两个任意线段的比,而且欧几里得还曾非常系统地加以发展。他们对那种比的运算,正像我们今天对于任意实数的运算一样。欧几里得的某些定义确实给了现代无理数理论以有力的启发。此外,对线段比所用的

术语也与任意自然数之比全然不同,后者称为 ἄϱιθμος,而线段之比,即任意实数,称为 λόγος。

我想对"irrational"("无理的")这个词再补充说上几句。这个词无疑是由希腊词"ἄλογος"译成拉丁文的,但是这个希腊词的大致意思是"不可表达的",意为这些新的数(或线段比)不能像有理数那样用两个整数之比来表达①。由拉丁词"ratio"("比")造成了一种误解,以为这个词只不过表示"reason"("理性")的意义,所以"irrational"(从词根上说,"irrational"本为"不可比"之意——中译者)就被误解成"unreasonable"("不合理的")的意义,致术语"irrational number"("无理数")至今似乎仍含有这一层意思。

无理数的一般概念首先出现在 17 世纪末,它是引入十进制小数的结果,而十进制小数的使用是随着对数表的出现而在那时确立下来的。如果我们把一个有理数变换成一个十进制小数,那么有时得到无尽小数,有时得到有尽小数,不过总是循环的。最简单的例子是 $\frac{1}{3}=0.333\cdots$,即小数点后立刻开始的循环节由一位数字组成的循环小数。这样,我们立刻就会想到其数字可根据任何定则而出现的非循环小数,这个非循环的小数凭直觉也就可以看做是一个确定的数,因而是非有理数。无理数概念的建立,就是这样。它在某种程度上是自然而然地产生的,是考虑十进制小数的结果。因此,从历史上来看,无理数的产生过程,是同我们已经讲过的负数产生过程一样的。计算的需要迫使引进新的概念,至于其本质或促成的因素,人们并不过多考虑。人们只是用以进行运算,特别是由于其十分有用,屡试不爽。

直到 19 世纪 60 年代,才感到有必要对无理数的基础确立一个较为精确的算术方式的表述。大致在那个时候,魏尔斯特拉斯在他所作的一个讲演中提到了这一点。1872 年点集理论的创始人,哈勒

① 见特罗麦克教科书,第 2 段,第 2 卷第 71 页。

的康托和布隆斯威格的戴德金,同时而又互相独立地为此奠定了一般基础。下面我想用几句话来解释一下戴德金的观点。假定我们具有关于有理数的全部知识,但排除一切空间概念,因为空间概念迫使我们立即联系到数序列的连续性观念。依据这样的理解,为了求得无理数的纯算术定义,戴德金在有理数域内建立了"分割"的概念。若 r 为任一有理数,它把全部有理数分成 A, B 两部分,致使 A 中的每一个数小于 B 中的任一数,而每一个有理数属于这两类数中的一类。A 是小于 r 的一切有理数的总体,B 是大于 r 的一切有理数的总体,r 本身可以任意看做属于这一类或另一类。除了这些正常"分割"外,还存在"非正常分割",即把一切有理数分成性质相同的两类,但不同之点在于它不是由某个有理数如上面的 r 那样造成的两类,即 B 类中不存在最小的有理数,A 类中也不存在最大的有理数。譬如说,$\sqrt{2} = 1.414\cdots$ 就是这种不正常分割的一个例子。实际上,每一个无尽十进制小数都已定义一个分割,若把大于此无尽小数的每个近似值的有理数归入 B 类,而把其他每一个有理数归入 A 类。这样,A 类中每一个数就等于或至少被一个近似数(因而也被无穷多近似数)超过。很容易证明,若为循环小数,则此分割即为正常分割,若为非循环小数,则此分割为非正常分割。

戴德金以这些考虑为基础建立了定义,他的这个定义从纯逻辑的观点去看只能当作一个任意的约定:有理数域内的一个分割因其为正常分割或非正常分割而称为有理数或无理数。由此立刻得出相等的定义:两个数若在有理数域内产生相同的分割则称为相等。根据这个定义,立即可以证明这样的一些例子,如 $\frac{1}{3}$ 等于无尽小数 $0.333\cdots$,如果我们接受这个观点,就必须要求有一个证明,即要求按给出的定义提出一个推理的过程。其实,如果采取朴素的方法,这个证明就可能会显得完全没有必要了。此外,这样一个证明也是直接的,因为我们联想小于 $\frac{1}{3}$ 的每一个有理数最终会被近似小数超过,而

近似小数又小于超过$\frac{1}{3}$的每一个有理数。魏尔斯特拉斯的讲演中给出如下相应的定义：若两个数之差(指较大者减去较小者)可小于任何预先给定的常数(不论此常数多小)，则此两数称为相等。这个定义同上述解释的联系是明显的。如果我们联想到 0.999…为什么等于1，那么后一定义就更使我们注目；此两数之差当然小于 0.1、小于 0.01等等，即按定义它们相等。

如果我们问怎么能把无理数纳入通常的数系而且同样地进行运算，那么回答是：单调性定律对于 4 种基本运算都是成立的，原则如下：如果想对无理数进行加法、乘法等运算，可以把每个无理数放在两个相距越来越近的有理数之间，然后对这些界限进行所期望的运算。由于单调性定律成立，因此运算结果也会在差不断缩小的数限以内。

我不必再把这些道理更详细地加以解释，因为在许多书里，特别是在威氏和伯氏的书里可以很容易地找到非常好读的论述。我希望除了我这里所推荐的这些书，你们能去读更多的书，以便更充分了解无理数的定义。

我在这里倒很想谈谈你们在上述书中未必能找到的东西，即在建立了这个算术理论以后怎样把它运用到其他领域里去。这特别是指解析几何，因为照朴素的观点，解析几何是无理数的起源(而从心理学上看也确实如此)。回想一下前面讲过的坐标轴，上面标有原点以及有理数点。据此可说上述理论的应用取决于下一个基本原则：对于每一个有理数或无理数，各存在一个以此数作为坐标的点；反过来，对于坐标轴线上的每一个点，也各存在一个有理数或无理数作为点的坐标。这个原则是一门知识所开始，随后一切都要由逻辑推导而来。但它本身不能得到逻辑证明，所以这个原则可以适当地称之为公理。这种公理从直觉上看是不言而喻的，或者可以按各人的天赋当作或多或少是任意的约定来接受。关于实数和直线点之间一一对应的公理，通常称为康托公理，因为康托是第一个明确提出这个公

理的人(在 1872 年《数学年刊》第五卷中)。

这里来谈谈关于空间观念的本质。空间观念的起源有两个,一个是对空间观念的直觉,可以通过量度而直接意识到。另一个就完全不同了,它是主观的理想化的直觉,也许可以说是我们与生俱来的空间观念,它超越感官观察的不精确性。在讨论数的观念时我曾向你们指出一个类似的区别,这种区别可以最恰当地刻画如下:谈到一个小的数,如 2,5 或者 7,我们马上就明白是什么意思,但是对于大一点的数,如 2 503,我们就没有那种直接的直觉了。这里,直接的直觉已被一种有序数列的主观直觉代替了,这种有序数列是我们通过数学归纳从开头几个数推导出来的。空间观念也有类似的情况。例如,如果我们来想两点之间的距离,那么我们只能以有限的精确程度估计它或测定它,因为我们的眼睛不能辨别长度差异小于一定限度的两个线段的不同。这就是直觉阈的概念,在心理学中有非常重要的作用。即使用最高精度的仪器来补助眼力之不足,本身仍然存在这种现象,因为有种种物理的特性使我们不能超越某种限度的精确性。例如,光学告诉我们,随着色彩的不同,光波波长的变化约为 $\frac{1}{1\,000}$ 毫米(=1 微米)数量级之小;还告诉我们大小为这种数量级的物体,即使用最好的显微镜也看不清楚,因为出现了衍射,所以任何光学的像也不能把细微处一一精确地再现出来。其结果是不能通过直接的光学手段测得小于 1 微米的长度,所以若以微米计所测的长度,则只有前三位小数有意义。同样地,在一切物理观察和测度中,我们都会遇到这样不可逾越的阈值,决定着以微米计算并表示的长度精确性的最高极限。超过这个极限的说法都没有意义,都是无知或欺人之谈。我们常常在矿泉水的广告中发现那种过于精确的数字,把实际上随时间变化的含盐比说到百分之零点零几,可这是不可能按重量测到那么精确的。

同精确性有限制的经验的空间观念相反,抽象的或观念化的空间观念都要求有不受限制的精确性。根据康托公理的观点,这正好

与我们对数的概念所下的算术定义相对应。

同我们把空间观念这样划分相协调，自然也就可以相应地把数学分成两部分，称之为近似数学及精确数学。如果我们打算用解方程 $f(x)=0$ 来说明这个区别，可以说，近似数学，正像经验的空间观念一样，并不关心 $f(x)$是否正好等于零，只关心其绝对值$|f(x)|$恒在可达到的精确性的阈值 ε 以下。符号 $f(x)=0$ 仅为不等式 $|f(x)|<\varepsilon$ 的简写，这才是我们真正关心的。只有在精确数学中，才坚持要不折不扣地满足 $f(x)=0$。由于在应用上只有近似数学才起作用，因此我们也许可以冒昧地说我们所需要的只是这个数学分支，而精确数学只是给孜孜以求的人得到智能上的快乐，同时对近似数学的发展提供可贵的、必不可少的支持。

为了回到正题上来，我再补充一句：无理数的概念当然只是精确数学范围内的概念，因为两点之差是几微米的一个无理数这种说法，没有任何意义。正如我们大家所看到的，当我们的尺是以米计的时候，六位以下的小数就失去了意义。因而在实践中，我们可以用有理数来代替无理数而不必担心。但是在结晶学中会讲到有理指数的定律，天文学中有两个行星的运转周期之比是有理数或无理数也必须区分，这好像和上面讲的情况有矛盾，但实际上这些表述方式只是表现了语言的多义性而已，他们用的有理及无理的含义与到这里为止我们所用的含义完全不同，即与近似数学里的含义完全不同。按照这种说法，两个量成有理数比，可能是指它们是两个小整数之比，如$\frac{3}{7}$，而同时可能称$\frac{2\,021}{7\,053}$为无理数比。我们一般说不清分子和分母究竟要有多大才属于第二种情况，因为这取决于具体问题。这一切有趣的关系，我在 1901 年的夏季讨论班上都讨论过了。那些讲稿后于 1902 年石印，将构成本书的第 3 卷（见第 3 版序），名为：《微积分在几何上的应用》，副标题为："原则的修订"（C · H · 穆勒协助完成）。（注：此中译本仅译出第 1 卷和第 2 卷，第 3 卷容后补译——译者）

最后，请允许我用三言两语谈谈要我到中学里教这些内容我会怎么教的问题。无理数的精确理论既未必适合大多数学生的兴趣，也超过他们的接受能力。一般来说，学生对有限精确性的结果已感到满足。对于精确度到$\frac{1}{1\,000}$毫米的结果，他们就会在赞许之中带着惊讶，而不会要求无限的精确性。对于普通程度的学生，只要通过例子一般讲明白无理数就足够了，平常也是这么做的。特别有天资的个别学生肯定会要求更完整的解释，给予这些学生以补充解释而不牺牲多数人的兴趣，在教师方面来说，就是值得赞扬的教学技巧了。

第三章

关于整数的特殊性质

我们现在就要翻开新的一章了，这一章所涉及的是当今的整数理论，即数论或所谓狭义的算术。我首先列出在中学课程中出现的、有关这门学科的个别的问题。

(1) 数论的第一个问题是可除性的问题：一个数是否可被另一个数所整除?

(2) 可以给出简单的规则，使我们能很容易判定能否用较小的数(2，3，4，5，9，10 等)去整除任何给定的数。

(3) 有无穷多的素数，即除去 1 和本身以外无其他整数因子的数：如 2，3，5，7，11 等。

(4) 如果我们知道给定整数的素因数分解，就掌握了该整数的一切性质。

(5) 在化有理分数为十进小数时，数论起着重要的作用；在数论中可以证明由有理数化成的十进小数为什么必定是循环的，以及循环节有多长。

尽管那样的问题可以放到中学里来考虑，教给 11 岁至 13 岁之间的学生，但只是到以后几年才会在零散的地方出现数论，顶多也只考虑以下各点。

(6) 偶尔会教连分数，而且不是所有的学校都学。

(7) 有时也教丢番图(Diophantus)方程，即带几个未知数的方程，其中未知数只取整数值。前面谈的毕达哥拉斯数，提供了一个例子；这里要处理 3 个整数，满足方程

$$a^2 + b^2 = c^2。$$

(8) 圆的等分问题同数论有密切的联系，不过这种联系在中学里从来也没有讲过。如果我们想把一个圆分成几个相等的部分(自然只用直尺和圆规来分)，那么对于 $n = 2, 3, 4, 5, 6$，是很容易办到的。但是对于 $n = 7$，就办不到了，所以当我们在中学里讲到这个问题时就只好免开尊口了。实际上，通常不是斩钉截铁地说，当 $n = 7$ 时这种作图真的不可能，因为这个事实的解释深深地植根于数论中。为了预先防止发生误解(不幸它是经常发生的)，请允许我强调一下，我们这里所涉及的，又是精确数学的一个问题，对于应用是没有意义的。在实践中，即使在可以得出"精确"作图法的情况下，通常也不会用它，因为在近似数学的领域，圆之可以任意等分，用简单而巧妙的实验方法更合适；而且任何规定的实际可能的精确度，都可以达到。每一个制造等分圆仪表的机械师，都是采取这样的方法的。

(9) 在中学课程中有一个地方接触到比较高深的数论，即求圆面积时计算 π。我们通常用某种方法确定 π 的前几位小数，或许也偶尔提到 π 的超越性的现代证明，这个证明一劳永逸地解决了用直尺和圆规来求圆面积的老问题。在这本书的结尾，我将详细讨论这个证明。暂时我只限于精确表述这样一个事实，即数 π 不满足带有整系数的任一代数方程

$$a\pi^n + b\pi^{n-1} + \cdots + k\pi + 1 = 0。$$

特别重要的是系数必须是整数，这个问题之所以属于数论的范畴，道理就在这里。当然，我们在这里又单纯涉及精确数学的问题，因为只是在这个意义上 π 的数论性质才有意义。近似数学只满足于确定前几位小数，使求出的圆面积达到所要求的精确度就可以了。

我已经对你们大致说明了数论在中学里的地位。现在让我来讨论一下它在大学教育及科学研究中的适当地位。就这一方面来说，我想根据对数论的态度把从事研究的数学家分成两类：一类我称为

满腔热情派；另一类是无所谓派。在前者看来，哪一种科学都没有数论那么完美而重要，都没有那么清晰而精密的证明，定理更没有严格到数论的定理那么无懈可击的程度。高斯(Gauss)说过："如果数学是科学中的女皇，那么数论就是数学中的女皇。"另一方面，抱无所谓态度的人却觉得数论离他们太远了；他们对数论的发展没有什么兴趣，简直是躲开它。大多数学生的态度就属于第二类。

这种泾渭分明的态度，其原因我想可以归结如下：一方面，数论对于一切比较透彻的数学研究都是基本的东西，即使从完全不同的领域出发，最后总是纷纷接触到相对简单的算术问题。另一方面，纯数论是极端抽象的，不是每一个人都有天赋来愉快地领会这样抽象的东西，大多数课本拼命把这个课程讲得无比抽象，结果反而使人对它望而生畏。我认为，如果结合一些图像和适当的数值，数论也许更能为大家接受，更能引起普遍的兴趣。尽管数论的定理在逻辑上并不需要借助于它们，但有了这些手段还是能帮助人们的理解的。我在1895—1896学年课程中作了这样的尝试，后来H·闵可夫斯基在《丢番图近似》一书中也采取了类似的作法。我的课是比较初步的引论性质的，但闵可夫斯基却是早期详尽探讨数论中各种专门问题的人。

至于教科书里讲的数论，你们往往可以在代数教科书里找到所需要的一切。在大量的实数理论书中，我要专门提一下巴赫曼(Bachman)的《新数论基础》。

在下面对数论比较专门的讨论中，我将注意上面讲到的那几点，特别努力尽可能用一些形象化的手段把它们讲清楚。我只限于提供对教师可贵的材料，丝毫也不打算把它写成可以马上拿来教学生的形式。我的阅卷经验告诉我必须这样做，因为我从考卷中发现，考生的数论知识往往不出几个术语的范围，术语背后的意义，他们并没有了解透彻。每个考生都会对我说，π是超越数，但是许多人并不了解这是什么意思；有一次我还听说，超越数既不是有理数，也不是无理

数。同样地，有的学生对我说素数是无穷的，但对于这么简单的证明，他竟然完全不知道。

我就从这个证明开始关于数论的讨论。我假定你们已经记住了这一节开始所列出问题的前两点。从历史渊源上说，我提醒你们注意，这个证明是欧几里得传下来的，他的《几何原本》不仅包含他的几何体系，而且也包含用几何语言写的代数和算术的知识。欧几里得传下来的关于存在无穷多素数的证明如下：假设素数的序列为有限的，设此序列为 2，3，5，…，P，因而数 $N=(1\cdot 2\cdot 3\cdot \cdots \cdot P)+1$（中译者注：英译误为 $N=(1\cdot 2\cdot \cdots \cdot P)-1$），不可以用 2，3，5，…，$P$ 中任何一个数来整除，因为总有余数 1。据此，N 本身必然是一个素数，或存在大于 P 的素数，两者必居其一，其中任一个皆与假设矛盾。证毕。

关于本节开始提的第四点，即整数之素因数分解问题，我提请你们注意一本老的素因数表：切纳克(Chernac)的 *Cribum Arithmeticum*[①]。这是一本值得称赞的大书，从历史观点来看，由于它十分可靠，因此更值得注意。表的名称使人想起埃拉托斯特尼(Eratosthenes)的筛。它所依据的思想是：我们应从所有整数的序列中逐一舍去可用 2，3，5，…整除的数，留下的就只有素数了。切纳克把 102 000 以下所有不可以用 2，3，5 来整除的整数都分解成素因数，而所有素因数都标上一横。上述范围以内的一切素数，是在切纳克的书中第一次给出的。19 世纪期间，找出了 9 百万以内的全部素数。

我现在转到第五点上，即化分数为十进小数问题。要知道完整的理论，建议你们参考韦伯—韦尔斯坦因的书，这里只通过一个典型的例子来说明其原则方法。考虑一分数 $\frac{1}{p}$，其中 p 为 2 和 5 以外的素数。由此可证明，$\frac{1}{p}$ 等于一无尽循环小数，循环位数 δ 是使得 10^{δ}

---

① 切纳克，Deventer，1811 年。

除以 p 余 1 的最小的指数 δ，或用数论的语言来说，δ 是满足下列同余式的最小指数

$$10^{\delta} \equiv 1 (\bmod p)。$$

为了证明这一结论，首先要求这个同余式恒有解。这一点利用费马(Fermat)小定理便可证明①。费马小定理说，对于 2 和 5 以外的每一个素数 p

$$10^{p-1} \equiv 1 (\bmod p)。$$

我们这里略去这个基本定理的证明，因为这个定理是每一个数学工作者的常规工具之一。其次，我们必须从数论中借用下一定理：这里所说的最小指数 δ 或为 $p-1$ 本身，或为 $p-1$ 的一个因数。我们可以把这个定理运用于给定值 p，得知 $\dfrac{10^{\delta}-1}{p}$ 是一个整数 N，故有

$$\frac{10^{\delta}}{p} = \frac{1}{p} + N。$$

如果我们现在考虑把$\dfrac{10^{\delta}}{p}$以及$\dfrac{1}{p}$转换为小数，那么两者的小数部分各位数字必相同，因为其差是一个整数。但是，既然$\dfrac{10^{\delta}}{p}$是从$\dfrac{1}{p}$的小数表示将小数点向右移动 δ 位而求出的，由此可见$\dfrac{1}{p}$的小数表示并不会因此而改变。换言之，$\dfrac{1}{p}$的十进小数表示是某 δ 位数字不断重复而构成的。

现在要知道不可能有更小的 δ' 位的循环，其中 $\delta' < \delta$。为此，只要证明每一个这样 δ' 必须满足同余式 $10^{\delta'} \equiv 1 (\bmod p)$ 即可，因为我们知道 δ 是这个同余式的最小解。利用前面的论证逆推过去，即可

① 费马定理有两个：一个是这里用的，称为费马小定理；另一个是著名的费马大定理。以后我都加上了大小两个字。——中译者

证明这一点。由我们的假设可知，$\frac{1}{p}$和$\frac{10^{\delta'}}{p}$的小数部分重合，因而$\frac{10^{\delta'}}{p}-\frac{1}{p}$是一个整数 N'，所以 $10^{\delta'}-1$ 可被 p 整除，或换言之，$10^{\delta'}\equiv 1 \pmod p$。证毕。

下面我向你们提供几个最简单的启发性例子。这些例子说明，δ 可以取很多不相同的值，包括小于和等于 $p-1$。首先注意，对于

$$\frac{1}{3}=0.333\cdots$$

循环节的位数是 1。事实上，$10^1\equiv 1 \pmod 3$。

同样地，得

$$\frac{1}{11}=0.090\,9\cdots,$$

由此 $\delta=2$，相应地，$10^1\equiv 10$，$10^2\equiv 1 \pmod{11}$。

最大值 $=p-1$ 出现在下一例子中

$$\frac{1}{7}=0.142\,857\,142\,857\cdots,$$

这里 $\delta=6$，实际上是有 $10^1\equiv 3$，$10^2\equiv 2$，$10^3\equiv 6$，$10^4\equiv 4$，$10^5\equiv 5$，$10^6\equiv 1 \pmod 7$。

现在让我们以相似的方式来研究第六点，即连分数。不过我不用通常的抽象算术的方式来讲，因为你们在别的书里也可以看到，如韦伯—韦尔斯坦因的书里就有。我将利用这个机会向你们说明，可以通过几何的图形把数论讲得很清楚、很容易理解。我们用这种手段来讲数论，其实只是步高斯和狄利克莱(P. Dirichlet)的后尘而已。把几何方法摒之于数论的门外，是后来的数学家干的，大约是从 1860 年以后。当然，我在这里所能讲的，只是最重要的思想线索和定理，但不给证明，因为我假定你们对于连分数的基本理论并不是完全的陌生。我有一本石印的数论讲义，里面讲得很透彻，同时可以参

阅我的《数学论文集》第二卷 209 页至 211 页。

你们知道给定正数 ω 是怎样展开成为连分数的。我从 ω 中分出最大的正整数 n_0，记作

$$\omega = n_0 + r_0,\text{其中 } 0 \leqslant r_0 < 1。$$

因而若 $r_0 \neq 0$，我们就像对待 ω 一样处理 $\frac{1}{r_0}$

$$\frac{1}{r_0} = n_1 + r_1,\text{其中 } 0 \leqslant r_1 < 1。$$

同样地继续写下去：

$$\frac{1}{r_1} = n_2 + r_2,\text{其中 } 0 \leqslant r_2 < 1,$$

$$\frac{1}{r_2} = n_3 + r_3,\text{其中 } 0 \leqslant r_3 < 1,$$

$$\cdots\cdots$$

如果 ω 是有理数，那么这个过程在有限步以后就终止，因为在那种情况下必会出现等于零的余数 r_v；否则这个过程就会无限地进行下去。在任一情况下，我们都把 ω 的连分数展开式记作

$$\omega = n_0 + \cfrac{1}{n_1 + \cfrac{1}{n_2 + \cfrac{1}{n_3 + \ddots}}}。$$

举例来说，π 的连分数是

$$\pi = 3.141\,592\,65\cdots = 3 + \cfrac{1}{7 + \cfrac{1}{15 + \cfrac{1}{1 + \cfrac{1}{292 + \ddots}}}}。$$

我们如果在第一个、第二个、第三个……偏分母之后停止展开，

那么就称之为收敛子的有理分数

$$n_0=\frac{p_0}{q_0},\ n_0+\frac{1}{n_1}=\frac{p_1}{q_1},\ n_0+\cfrac{1}{n_1+\cfrac{1}{n_2}}=\frac{p_2}{q_2},\ \cdots。$$

这些分数给出的近似数是非常接近数 ω 的，或说得更精确一点，其中每一个分数给出一个近似数，比其他任何一个没有较大分母的有理数更接近数 ω。由于这种性质，连分数有重大的实用价值。当我们希望用一个具有小分母的分数去最佳逼近一个无理数，或是去最佳逼近一个具有较大分母的分数时，就用得着连分式。从下列 π 的连分数的一些收敛子化成的小分数，可以使我们看到这些近似数是怎样接近于 $\pi=3.141\,592\,65\cdots$ 的

$$\frac{p_0}{q_0}=3,\ \frac{p_1}{q_1}=\frac{22}{7}=3.142\,85\cdots,$$

$$\frac{p_2}{q_2}=\frac{333}{106}=3.141\,509\cdots,\ \frac{p_3}{q_3}=\frac{355}{113}=3.141\,592\,292\cdots。$$

在这个例子中你还会观察到，收敛子交替地小于和大于 π。大家知道，这一点具有普遍性，即 ω 的连分数相继的收敛子交替地小于和大于 ω，并把 ω 夹在不断缩小的数限之间。

现在让我们用几何图形把这些想法形象地表示出来。我们限于考虑正数，把 xy 平面第一象限中的一切有整坐标的点标出来（见图3.1），形成所谓格点。让我们来看这些格点。从 O 到点 $(x=a, y=b)$ 的向径，其方程为

$$\frac{x}{y}=\frac{a}{b}。$$

相反地，在每一条这样的射线：$\frac{x}{y}=\lambda$（其中 $\lambda=\frac{a}{b}$ 是有理数）上有无穷多的整点 (ma, mb)，其中 m 是任意正整数。从 O 向外看，我们在一切有理方向上，且仅在这样的方向上看到格点。我们的视野中

"处处稠密地"但不是完全地及连续地充满"星点",这与我们仰观银河的景象相像。在无理射线 $\frac{x}{y}=\omega$(ω 为无理数)上除了 O 本身以外没有一个整点,这是非常令人惊讶的。如果我们回忆戴德金的无理数定义,就会明白那样一个射线形成了整点场上的一个分割,把点分成两个点集,一个在射线之右,一个在左。如果问这些点集怎样向射线 $\frac{x}{y}=\omega$ 收敛,那么就可以发现,这与 ω 的连分数有着非常简单的关联。标出对应于收敛子$\frac{p_\nu}{q_\nu}$的每一个点 $(x=p_\nu,\ y=q_\nu)$,可以看到,过这些点的射线,交替地从左和从右逐渐更加近似于射线 $\frac{x}{y}=\omega$,正像数 $\frac{p_\nu}{q_\nu}$ 近似于数 ω 一样。此外,如果利用 p_ν, q_ν 的已知数论性质,那么就得出以下定理:想象在一切整点上钉上木钉或铁针,用两根细绳从 ω 射线出发分别向左右拉紧,如此围绕两个整点集拉成的两个绳索凸多边形的顶点,恰好是整点$(p_\nu,\ q_\nu)$,这些点的坐标是收敛子的分子和分母,它们相继收敛于 ω,左边多边形的诸顶点是那些偶序收敛子,右边多边形的诸顶点是那些奇序收敛子。这样就得出了一个新的、不妨说是极为图形化的连分数的定义。

图 3.1 所示对应于下面的例子:

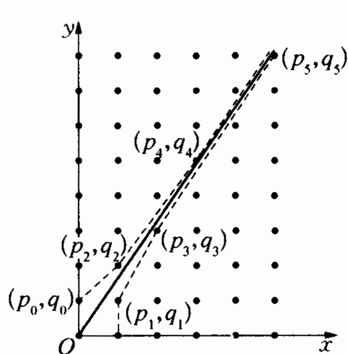


图 3.1

$$\omega=\frac{\sqrt{5}-1}{2}=\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\ddots}}}}。$$

这就是同正十边形相联系的无理性。在这个例子中,两个多边形的最初几个顶点是

左:$p_0=0$, $q_0=1$; $p_2=1$, $q_2=2$; $p_4=3$, $q_4=5$; …;

右：$p_1=1$，$q_1=1$；$p_3=2$，$q_3=3$；$p_5=5$，$q_5=8$；…。对于 π 的值 p_ν，q_ν 就增加的快得多，以致很难画出相应的图示。

我现在来研究第七点，即毕达哥拉斯数。我们这里将利用空间概念，但利用的方式有些不同，不考虑我们想要求其整数解的下列方程：

$$a^2+b^2=c^2, \tag{1}$$

我们置

$$\frac{a}{c}=\xi,\ \frac{b}{c}=\eta, \tag{2}$$

而考虑方程

$$\xi^2+\eta^2=1, \tag{3}$$

并求满足这个方程的一切有理数对 ξ，η。我们设想在 $\xi\eta$ 平面上标出一切有理点 ξ，η(即带有理坐标 ξ，η 的点)，这些点稠密地分布于 $\xi\eta$ 平面上。$\xi^2+\eta^2=1$ 是这个平面上以原点为中心的单位圆。我们的问题是考察这个圆怎样穿过有理点的稠密集，特别是看它包含哪些有理点。我们早已知道几个这样的点，如它与坐标轴的交点，考虑其中一个点 S ($\xi=-1$，$\eta=0$)(图3.2)。通过 S 的一切射线由方程(4)给出

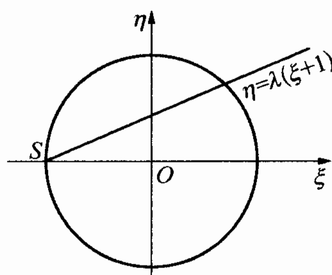


图 3.2

$$\eta=\lambda(\xi+1)。 \tag{4}$$

根据参数 λ 是否为有理数，我们称这样的射线为有理的或无理的。现在我们得到一个正逆均成立的定理：从 S 出发穿过圆的每一个有理点的射线是有理射线；每一条有理射线与圆相交于一个有理点。定理的前半部分是很明显的。我们将(4)式代入(3)式，以证明后半部分。于是得到交点横坐标应满足的方程：

$$\xi^2+\lambda^2(\xi+1)^2=1,$$

或
$$(1+\lambda^2)\xi^2+2\lambda^2\xi+\lambda^2-1=0。$$

我们知道这个方程的一种解是 $\xi=-1$，对应于交点 S；对于另一个解，通过简单计算即得

$$\xi=\frac{1-\lambda^2}{1+\lambda^2}。\tag{5a}$$

从(4)式得对应的 η 坐标是

$$\eta=\frac{2\lambda}{1+\lambda^2}。\tag{5b}$$

由(5a)式及(5b)式可知，若 λ 是有理的，则第二个交点是有理点。

我们的正逆都对的定理已经完全证明，还可以把它叙述如下：圆的一切有理点都由公式(5)来表示，其中 λ 为一任意的有理数。这就解决了我们的问题，只需要转换成整数就可以了。为此置

$$\lambda=\frac{n}{m},$$

其中 n, m 为整数，由(5)式得

$$\xi=\frac{m^2-n^2}{m^2+n^2}\quad\eta=\frac{2mn}{m^2+n^2},$$

表示出(3)式的一切有理数解的总体。因而，原方程(1)的一切整数解，即一切毕达哥拉斯数，可由下列方程给出

$$a=m^2-n^2,\ b=2mn,\ c=m^2+n^2,$$

并且若 m 和 n 取一切互素整数对，则得没有公约数的全部解。这样，平常看来非常抽象的一个结果，从图形上得到了推导。

就此我想来讨论费马大定理。正是紧紧模仿古代几何学家的作法，我们来推广毕达哥拉斯数问题，用以下方式把它们从平面推广到三维及三维以上空间。我们问：两个正整数的立方和仍旧是整数的

立方吗？两个正整数的四次方之和还是正整数的四次方吗？等等。一般地，设 n 为任意正整数，方程

$$x^n + y^n = z^n$$

有正整数解吗？对于这个问题，费马在以他命名的大定理中回答说：除 $n = 1$ 和 $n = 2$ 的情形而外，对于任意的正整数 n，方程 $x^n + y^n = z^n$ 没有正整数解。

现在我来插入几句历史。费马生于 1601—1665 年间，是当时图鲁兹地方的议会评议员，也就是一个法律学家。但他对数学涉猎很广，而且在数学方面取得的成就也最大，所以他可以算是最伟大的数学家之一。在解析几何、微积分和概率论的奠基人之中，费马的名字占有一个永恒的地位。但特别重要的是他在数论方面的成就。他在这个领域取得的一切研究成果，都是以他在一本丢番图的书的书沿上的批注出现的。丢番图也是古代的一个著名的数论大师，生活在亚历山大，大约是公元 300 年间的人，比欧几里得晚 600 年。费马的书是在他死后 5 年由他的儿子发表的，保持着原来批注的形式。费马生前自己没有发表过任何作品，但他与最著名的同代人有大量的通信，所以他的发现早已为人所知，尽管只有其中一部分。我们现在讨论的这个著名定理，就是出现在那本丢番图的书的批注版中。关于那个定理，费马曾在信中写到过。他说他已经找到了一个非常妙的证明，但书眉太窄了，写不下[①]。可是，时至今日，还没有一个人找到他这个定理的证明[②]！

为了使我们对费马大定理的要点有所了解，像在 $n = 2$ 的情况下一样，我们首先来问方程

$$\xi^n + \eta^n = 1$$

的有理解是否存在，即问表示这个方程的曲线与 $\xi\eta$ 平面上的全部有

---

① 见巴黎科学院版本：*Œuvres de Fermat*，第 1 卷，第 291 页，巴黎，1891 年，以及第 3 卷，第 241 页，巴黎，1896 年。

② 这一猜想已于 1994 年被怀尔斯(A. Wiles)解决。——中译者

理点的关系。对于 $n=3$ 及 $n=4$，曲线形状大致如图 3.3 及图 3.4 所示。当 $n=3$ 时，曲线中至少包含点 $\xi=0$，$\eta=1$ 及 $\xi=1$，$\eta=0$；当 $n=4$ 时，至少包含点 $\xi=0$，$\eta=\pm1$ 及 $\xi=\pm1$，$\eta=0$。费马的断定意味着，这些曲线与前面讲到的圆不同，它们穿过稠密的有理点集，但除了刚才指出的那些点以外，不通过另外任何一个有理点。

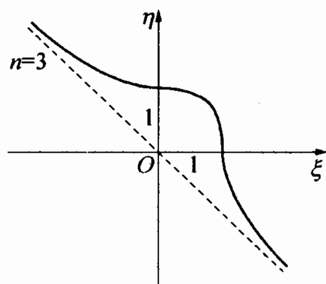


图 3.3

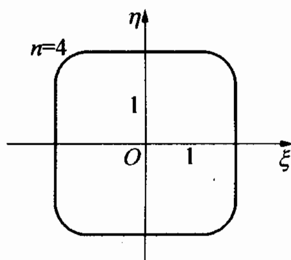


图 3.4

这个定理的有趣在于至今没有得到完全的证明，一切努力都是徒劳无功。在试图求证的人中，首先应该提到库默尔(E. Kummer)，他大大地推进了这个问题，把它同代数数论发生了联系，特别同 n 次单位根(割圆数)联系了起来。利用 1 的 n 次根，$\varepsilon=e^{\frac{2i\pi}{n}}$，我们确能把 z^n-y^n 分成 n 个线性因子，把费马方程记作下列形式

$$x^n=(z-y)(z-\varepsilon y)(z-\varepsilon^2 y)\cdots(z-\varepsilon^{n-1} y),$$

问题因而归结为按所述方式把整数 x 的 n 次幂分成由两个整数 z 和 y 及数 ε 构成的 n 个线性因子。库默尔对于这样的数所建立起的理论，同大家早已了解的、对于普通整数的理论非常相像，靠的是整除性和因数分解的概念。正因为如此，遂有整代数数的说法；特别说来，此处由于 ε 与圆的分割的关系，就有所谓分圆数。因此对于库默尔来说，费马大定理就是代数分圆数范围里的因子分解定理。他企图根据这个分析来推导定理的证明，实际上他也取得了一些成功，把 n 推导到了一个很大的数值，推导到了 100 以下的一切 n 值。但是

在大数之中出现了例外值，对于这些例外值还没有找到定理的证明，无论是他本人以及继起的数学家都没有做到这一点。

我的意见就讲到这里为止。关于这个问题的现状，以及关于库默尔的文章，你们可以在《数学百科全书》第 1 卷，第 714 页[1]希尔伯特的文章“Theorie der Algebraische Zahlkorper”的结尾部分找到详细介绍[2]。希尔伯特本人也属于继续并推进库默尔研究者之行列。

费马自称的“极妙的证明”，到底是不是沿这个方向，也确实难以作出断定。不过他未必会用代数数来进行推理，因为当时大家对虚数的意义还没有肯定。而且当时数论的研究还不发达。数论正是通过费马本人的手而大大得到推进。另一方面，我们也不能假定像费马这样的第一流数学家证明有误，尽管最伟大的数学家也犯过那样的错误。因此我们只好认为，他已经非常幸运地利用一个简单的想法求得了证明。不过由于我们对他探索的方向一无线索，因此可能只有期望通过系统地推广库默尔的结果来求得费马大定理的完全的证明。

可是在我们哥廷根科学会悬赏 10 万马克以求证费马大定理之后，这些问题又有了一层新的意义。这些悬赏基金是 1906 年逝世的数学家华尔夫斯凯尔（Wolfskehl）捐赠的。他大约毕生从事费马大定理的研究。死后从他的一大部分遗产中提出这笔数目以奖赏能确定费马大定理为真或能以简单的例子揭示其为不真的幸运儿。要反证其为不真当然也不是简单的事，因为这个定理已经证到 100 以下的指数，要重起炉灶，就得从非常大的数算起。

从我前面讲的可以明白，在了解情况，在了解库默尔及其后继人为证这个定理而付出多大努力的数学家看来，要获得这个奖金是多

---

① 见 P · 巴赫曼的文章《费马问题》，柏林，1919 年。该文为关于费马大定理的初步研究的综合报导。

② 费马大定理曾悬赏求证（早已失效），详情公布于 1908 年 Nachrichten d. Ges. d. Wessenshaften Zu Göttingen 通告，见该刊第 103 页及以后各页。详情也曾由其他许多数学杂志转载（例如，德国《数学年刊》，第 66 卷，第 143 页，《数学月刊》，第 134 卷，第 313 页）。

么困难。但是广大公众的想法不同，自从 1907 年夏天报上公布了悬赏的消息后（附带提一句，报纸并未被授权发表这条消息），我们收到了一大堆自称的证明。各行各业的人，包括工程师、中学教师、教士，还有一个银行家、许多妇女，都参加了投稿。这些人有一个共同点，就是对这个问题的严肃的数学意义毫无了解。不但如此，他们也没有努力去掌握有关这方面的资料，凭灵机一动就认为自己取得了答案，结果他们的工作不可避免地成为胡说八道。在 *Archiv für Mathematik und Physik*① 中，保存着 A·弗莱克(A. Fleck)(他的职业是开业内科医生)、曼恩钦博士(Ph. Maennchen)以及 O·佩龙(O. Perron)对那类证明的大量分析批评，如果读到这些文章，你就可以看到那类证明出现了什么样的谬误。这种胡闯乱杀的情况，读来使人感到可笑，同时又感到可悲，悲其无必要而已。我想提一个例子，这个例子是同我们所讨论的 $x^2+y^2=z^2$ 有关的。解题的人想对函数 $x^n+y^n=z^n$ $(n>2)$ 找出一个有理的参数表示，但这与 $n=2$ 的情况是不同的，从代数函数论去看，我们早已知道这个结果是不可能的。这个人忽略了这样一个事实：对于自变量的个别有理数，非有理函数可以取有理数值，因而自以为已证明了费马大定理。

对费马大定理就谈到这里为止，下面来谈第八个问题——圆的分割问题。我这里假定你们已了解复数，所以要用到复数 $x+\mathrm{i}y$ 运算，尽管我以后还要系统地讨论复数。这里说的问题其实就是圆的 n 等分或作一 n 边的正多边形问题。我们把圆定为 xy 复平面以原点为中心的单位圆，并取 $x+\mathrm{i}y=1$ 为 n 个分割点的第一个点(图 3.5，其中 n 等于 5)。因而 n 个顶点是

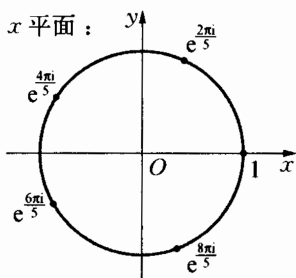


图 3.5

① 第 14—17 卷，第 18 卷(1901—1911 年)。

n 个复数

$$z = x + \mathrm{i}y = \cos\frac{2k\pi}{n} + \mathrm{i}\sin\frac{2k\pi}{n} = \mathrm{e}^{\frac{2k\pi\mathrm{i}}{n}}\ (k = 0,\ 1,\ \cdots,\ n-1)。$$

根据棣·莫弗(De Moivre)定理,它们满足方程

$$z^n = 1,$$

由此圆的分割问题变成为解这个简单的代数方程问题。既然这个方程中有有理数根 $z = 1$, $z^n - 1$ 可为 $z - 1$ 整除,则余下的 $n-1$ 个根应满足所谓分圆方程

$$z^{n-1} + z^{n-2} + \cdots + z^2 + z + 1 = 0,$$

即其系数皆为$+1$ 的 $n-1$ 次方程。

自古以来,对于可以用尺规(直尺和圆规)作什么样的正多边形问题就很有兴趣。古人已经知道,对于数 $n = 2^h$, 3, 5 (h 为任意整数),可以作正多边形,对于合数 $n = 2^h \cdot 3 \cdot 5$ 也同样可以。问题就到此为止,直到 18 世纪末才由年轻的高斯完全解决。他发现可以用尺规对形如 $p = 2^{2^\mu} + 1$ 的一切素数作出正 p 边形,但对其他素数都不可能。对于最初几个数值 $\mu = 0, 1, 2, 3, 4$, 据此公式实际上可得出素数

$$3,\ 5,\ 17,\ 257,\ 65\,537,$$

其中前两种情形是已经知道的,其他各项是新的。在这些情形中,正十七边形是特别著名的。首先发现用尺规作正十七边形的作图法的是高斯。另外,并不知道对于什么样的 μ 值,上述公式表达的 p 正好是素数。从欧拉那时就知道,对于 $\mu = 5$ 是一个合数。详细情况我就不再谈了,我只是想说明一般情况和这个发现的重要性。你们要知道有关正十七边形的详细说明,可以在韦伯—韦尔斯坦因的书中找到。

我想提请你们特别注意重刊在《数学年刊》(1903 年)第 57 卷和

《高斯著作集》第 10 卷第 1 页(1917 年)中的高斯日记。这是一本不起眼的小本子,高斯从满 19 岁前不久的 1796 年记起。第一则日记就涉及能不能作十七边多边形问题(1796 年 3 月 30 日)。高斯后来决定献身于数学,也正是由于这个早期的重要发现。对于每一个数学工作者来说,浏览这一本日记都很有意义,它可以使我们进一步紧紧地抓住高斯在数论、椭圆函数等领域所作出的奠基性发现的来龙去脉。

高斯的第一个伟大的发现是在他的老师和监护人霍法拉特·齐末尔曼(Hofrat Zimmermann)的鼓动下发表的,1796 年 6 月 1 日以短讯形式发表于《耶拿文学报》上,附了他老师写的一个短注[①]。高斯后来在他的奠基性的数论著作——*Disquisitiones Arithmeticae* (1801 年)中发表了证明。在这里,我们第一次发现这个定理的否定方面的证明。而这在他的短讯中是没有的,即:对于形如 $2^{2^\mu}+1$ 以外的素数,如对于 $p=7$,都不可能用尺规作图。这个不可能性证明是非常重要的,我要在这里向你们提出一个例子,因为广大公众对这类证明缺乏理解,我就更加愿意这样做了。要知道现代数学家已借助于这类不可能性证明,解决了整整一系列自古以来许多数学家虚掷时光所不曾解决的著名问题。除了正七边形的作图以外,我只提一下用尺规来三等分一个角和化圆为方的问题。但是现在还有许许多多没有一点高等数学知识,甚至对不可能性证明的性质也不懂的人,却要去解这些问题。他们的知识多半不超过初等几何的水平,他们通常这样做一个辅助线、那样作一个辅助圆,左试右试,弄得繁复无比,如果不用很多很多时间,世上没有一个凡人能走出迷宫,向他们指出作图中的错误。你叫他们去参考不可能性的数学证明也没有用,因为只有直接考虑他们本身"证明"、直接指出其证明中的错误,他们才肯信服。稍有名气的数学家每年都要收到一大堆那种请托的

① 也印在《数学年刊》,第 52 卷,第 6 页(1903 年),以及《高斯著作集》,第 10 卷,第 1 页(1917 年)。

东西，你们到了工作岗位上也会收到那种证明，你们最好预先做好思想准备，知道如何坚持自己的观点。所以，如果你们能深入浅出地掌握不可能性证明，那将是一件很好的事。

据此，我想向你们提出一个详细的证明，证明从精密几何的意义上说不可能用尺规作出正七边形来。大家都很了解，尺规作图都有等价的算术表示，表示为求一连串的平方根；反之，每一个那样的平方根的求法也可以用直线和圆相交的几何方式来表示。这一点，你们可以很容易地自己加以验证。于是，我们可以用解析方式来表述我们的论断：表示正七边形的六次方程

$$z^6+z^5+z^4+z^3+z^2+z+1=0$$

不能用有限的逐次平方根来求解。上述这个方程就是所谓反商方程，即对于每一个根 z，还有一个根$\frac{1}{z}$。如果我们把这个方程记作

$$z^3+z^2+z+1+\frac{1}{z}+\frac{1}{z^2}+\frac{1}{z^3}=0 \tag{6}$$[①]

那就显而易见了。如果我们取

$$z+\frac{1}{z}=x$$

为一个新的未知数，那么就可以通过把上述方程次数降低一半的方法进行简化。经简单计算，求得 x 的三次方程

$$x^3+x^2-2x-1=0, \tag{7}$$

于是立刻可以看出，方程(6)和方程(7)都可以或都不可以用平方根求解。此外，我们还可以结合七边形的几何构造把 x 表示出来。因为如果考虑复平面上的单位圆，就很容易看出下述关系是显而易见

① 为统一体例，书中的公式按节编号，下同。——编辑

的。如果把正七边形的圆心角记作 $\varphi=\frac{2\pi}{7}$，记住 $z=\cos\varphi+\mathrm{i}\sin\varphi$ 和 $\frac{1}{z}=\cos\varphi-\mathrm{i}\sin\varphi$ 是七边形的最接近于 $x=1$ 的两个顶点，那么 $x=z+\frac{1}{z}=2\cos\varphi$（图 3.6）。因此知道 x 就立刻可以作出正七边形。

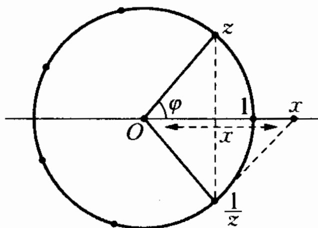


图 3.6

现在我们必须证明三次方程(7)不能用平方根来解。这个证明分成算术证明和代数证明两部分。我们第一步先证明方程(7)不可约，即其左边不能被分成系数为有理数的两个因子。如果这个方程可以约化，那么其左边必有一个带有理系数的线性因子，因而必有一个有理数 $\frac{p}{q}$（其中 p 和 q 是没有公因子的整数），使它为 0。即

$$p^3+p^2q-2pq^2-q^3=0。$$

因此 p^3 可被 q 整除，从而 p 就可以被 q 整除。同理，q^3 及 q 也必可被 p 整除。因此，$p=\pm q$，从而方程(7)必有根 $x=\pm 1$。但检验结果并不如此。

证明的第二部分就是要证明带有理系数的、不可约的三次方程不能用平方根来解。这本质上是代数性质的证法，但是由于它和算术证法有联系，我就在这里给出。让我们用正面断定的方式来说。如果带有理系数 A,B,C 的一个三次方程

$$f(x)=x^3+Ax^2+Bx+C=0 \tag{8}$$

能够用平方根来解，那么这个方程必有一个有理根，即为可约。因为存在有理根 α，等价于 $f(x)$ 存在有理因式 $x-\alpha$，因而可约。最重要的是：在给出这个证明之前，需要把一切可以用平方根来构成的式

子分类。更精确地讲,是把一切可以用有限个平方根和有限个有理数构成的式子分类。那样一个数的具体例子是

$$\alpha=\frac{\sqrt{a+\sqrt{b}}+\sqrt{c}}{\sqrt{d+\sqrt{e+\sqrt{f}}}},$$

其中 a, b, …, f 是有理数。当然我们谈的只是不能用有理数开出的平方根的情况。至于其他的一切平方根,则必须简化。那样的每一个表示式都是若干平方根的有理函数。我们的例子中有 3 个平方根。我们先考虑那种单独的平方根,不过它的被开方数可以具有要多复杂就多复杂的形式。其"阶"是指其中出现的层层相叠的根号的最大个数。在上一个例子中,α 的分子中分别有二阶和一阶根式,而分母的根式为三阶。

就一般的平方根表示式而言,我们考虑刚刚讨论过的各个"简单平方根表示式的阶数",一般的平方根表示式就是由它们经有理运算而得到的,其最大阶 μ 就称为这个表达式的阶数。在我们所举的例子中,$\mu=3$, 然而可能出现阶数同为 μ 的 n 个简单平方根表示式,所以它的第二个特征就是 μ 阶项的"项数",可以记作 n。这个项数可以看做是非常确定的,以致 n 个阶为 μ 的简单表示式中,没有一个可以用阶数等于或小于 μ 的其他简单表示式来有理表示。例如,阶为 1 的表示式

$$\sqrt{2}+\sqrt{3}+\sqrt{6},$$

其项数是 2 而不是 3,因为 $\sqrt{6}=\sqrt{2}\cdot\sqrt{3}$。对于上面给出的例子 α, $n=1$。

这样,我们在每一个平方根表示式上加上了两个有理数 μ, n, 我们用符号 $(\mu,\ n)$ 把它们结合起来,称为平方根表示式的"特征"或"秩",并且规定当两个平方根表示式的阶不同时,阶数低的表示式,其秩也低;当阶相同时,秩的高低由项数的多少来决定。

现在假定三次方程(8)的根 x_1 可以用平方根表示,说得清楚一点,可以用秩为$(\mu,\ n)$的表示式来表示。选取阶为 μ 的 n 项之一 $\sqrt{R}$,令 x 记作

$$x_1=\frac{\alpha+\beta\sqrt{R}}{\gamma+\delta\sqrt{R}},$$

其中 $\alpha,\ \beta,\ \gamma,\ \delta$ 至多含有 $n-1$ 个 μ 阶的项,R 的阶为 $\mu-1$。这里,$\gamma-\delta\sqrt{R}$ 当然不等于 0,因为 $\gamma-\delta\sqrt{R}=0$ 不是 $\delta=\gamma=0$(这是不可能的)就是 $\sqrt{R}=\gamma:\delta$, 即$\sqrt{R}$可借助于出现在 x 中的其他 $(n-1)$ 个 μ 阶项来作有理表示,因而$\sqrt{R}$是多余的。把分子、分母乘以 $\gamma-\delta\sqrt{R}$,得

$$x_1=\frac{(\alpha+\beta\sqrt{R})(\gamma-\delta\sqrt{R})}{\gamma^2-\delta^2\cdot R}=P+Q\sqrt{R},$$

其中 $P,\ Q$ 是 $\alpha,\ \beta,\ \gamma,\ \delta$ 的有理函数,即至多含有 $(n-1)$ 个 μ 阶项,此外,其他项的阶数只能是 $\mu-1$,所以它们至多有秩 $(\mu,\ n-1)$。把(8)式中的这个 x 值代入,得

$$f(x_1)=(P+Q\sqrt{R})^3+A(P+Q\sqrt{R})^2$$
$$+B(P+Q\sqrt{R})+C=0。$$

去掉括号,得一关系,形如

$$f(x_1)=M+N\sqrt{R}=0,$$

其中 $M,\ N$ 是 $P,\ Q,\ R$ 中的多项式,即 $\alpha,\ \beta,\ \gamma,\ \delta,\ R$ 的有理函数。如果 $N\neq0$,我们就应有 $\sqrt{R}=-\dfrac{M}{N}$,即$\sqrt{R}$可用 $\alpha,\ \beta,\ \gamma,\ \delta,\ R$,即用其他 $(n-1)$ 个 μ 阶项及低阶的其他项作有理表示。但根据假设,这是不可能的,正如上面所指出的。因而必然得出 $N=0$,因而 $M=0$。由此可以得出结论

$$x_2 = P - Q\sqrt{R}$$

也是三次方程(8)的一个根,因为同前面的方程一比较,立刻可以得出

$$f(x_2) = M - N\sqrt{R} = 0。$$

现在这个证明可以很简单地、令人吃惊地结束,如果 x_3 是上述三次方程的第三个根,则得

$$x_1 + x_2 + x_3 = -A,$$

因而

$$x_3 = -A - (x_1 + x_2) = -A - 2P$$

与 P 同秩,从而秩低于 x_1。

如果 x_3 本身是有理的,那么我们的定理就得到了证明;如果本身不是,那么我们就可以把它当作同样的一系列推理的出发点。所以其他的根之秩均更高只是一个幻想,特别来说,其中的一个根,其秩确实应低于 x_3 之秩。如果我们在各个根之间翻来覆去地这样做,那么每一次总是看到秩确实比我们想得更低,因而最终不得不得出阶数 $\mu = 0$ 的根。这表明三次方程中存在着一个有理根。我们不能再这样做下去了。因此,其他的两个根或本身是有理的,或者可以写为 $P = Q\sqrt{R}$,其中 P, Q, R 为有理数。由此表明,$f(x)$可分成一个二次因子和一个一次有理因子,因而是可约的。每一个不可约的三次方程,特别是我们所讨论的正七边形方程,是不可能用平方根求解的。因而证毕:正七边形不能用尺规作图。

你们看,这个证明的推导是多么简单、多么一目了然,它所需要的预备知识又是多么的少。尽管如此,某些步骤,特别是对于平方根表示式分类的解释,要求有一定程度的数学抽象能力。这个证明是否简单得足以说服上述的数学门外汉,使他们认识到打算用初等几何来证明是徒劳无功的,我不敢冒昧断定。不过对于那样的人,我们

应该慢慢地把证明解释清楚。

在结束的时候，我要提一些有关正多边形问题的文献，并连带提一下这里已涉及的一般几何作图问题上的文献。首先又要提到韦伯—韦尔斯坦因的书（第 18 版第 1 卷，第 17、第 18 两节）。其次请允许我提到我在 1895 年为哥廷根大学教师的一次集会而准备的纪念性小册子 *Vorträge über ausgewählte Fragen der Elementargeometrie*[①]。要了解更为详尽的内容，还可以提一下意大利波洛尼亚市 F · 恩里克斯（Enriques）编的《初等几何问题》，此书有德文译本，名为 *Fragen der Elementargeometrie*[②]，你们可以从中找到一切有关问题的资料，以代替我那本已出版的小册子。

现在我就结束数论的讨论，把最后一点——π 的超越性——留到这个课程的结尾来谈，下一章就转到数的概念的最后扩张上去。

① F · 塔格特（F. Tägert）整理，莱比锡，1895 年。此书有贝曼（Bemann）及史密斯（Smith）合译的英译本，名为《著名的几何问题》，吉恩（Ginn）出版公司，后由斯特切特（Stechert）于纽约重印。

② 由 H · 弗莱谢尔（H. Fleischer）译成德文，莱比锡，1907 年（第二版，1923 年）。请参阅德文第二部分"Die geometrischen Aufgaben, ihre Lösung und Lösbarkeit"。还可参阅 J · W · A · 杨的专题论文：《现代数学中的问题》。

第四章
复　　数

§4.1 通常的复数

让我以介绍某些历史事实作为开端。据说第一个用虚数的人是卡丹(G. Cardan),实际上是他在1545年解三次方程时偶然用到的。至于后来的发展情况,就像负数的情况一样,它之所以进入算术计算,开始并未得到一致赞同,甚至是违反个别数学家的意愿的;只是随着它被证明为有用,才逐渐获得了越来越广泛的使用。同时,数学家也对它并不完全满意。长期以来,虚数保持着某种神秘的色彩,正如今天每一个学生第一次听到奇怪的 $\mathrm{i}=\sqrt{-1}$ 时一样。我可以举出莱布尼兹1702年说的一句很重要的话作为证据。莱氏说:“虚数是圣灵的完美而奇妙的避难所,也差不多是介于存在和不存在之间的两栖类。”18世纪时,有关的概念还远没有澄清,尽管欧拉首先认识到了虚数对于函数论的根本意义。1748年,欧拉建立了下述奇妙的关系:

$$\mathrm{e}^{\mathrm{i}x}=\cos x+\mathrm{i}\sin x。$$

借此我们认识了初等分析中出现的种种函数之间的关系。到了19世纪,才最终对复数的性质有了清楚地理解。这里首先必须强调指出对复数所作的几何解释,因为大约在19世纪末,各种研究者都被引到这条路上来了。我只消提一个人的名字就可以了,这就是高斯。高斯无疑是对这件事的本质有最深入了解,并对公众发生过最持久影响的人。上面提到过他的日记毋庸置疑地证明,他在1797年已经完全掌握了这个解释,尽管很久以后才发表。19世纪的第二个成就

是为复数创立了纯形式的基础，使复数的讨论依赖于实数。这是 19 世纪 30 年代英国数学家首创的，这里略去细节不谈，不过你们可以在前面提到过的亨克尔的书中找到详细材料。

现在我来解释前面讲的两个最常用的基本方法。我们先站在纯形式的立场上来谈，把各个运算规则之间的相容性，而不是把对象本身的意义作为概念正确性的保证。据此观点，复数的引入，采取下述方式，完全排除了神秘的味道。

(1) 复数 $x+\mathrm{i}y$ 是 x，y 这两个实数的结合，即一个数对。关于这些数对，作以下的规定。

(2) 两个复数 $x+\mathrm{i}y$，$x'+\mathrm{i}y'$ 在 $x=x'$，$y=y'$ 时称为相等。

(3) 加法和减法根据下列关系来定义

$$(x+\mathrm{i}y)\pm(x'+\mathrm{i}y')=(x\pm x')+\mathrm{i}(y\pm y')。$$

一切加法规则由此导出，这是不难证实的。只有单调性法则按其原来的形式失去了有效性，因为复数本质上不像自然数和实数可以按其大小而有次序关系。为了简明扼要，我不讨论单调性法则的修正形式。

(4) 我们约定：除了始终认为 $\mathrm{i}^2=-1$ 以外，复数的乘法运算就像一般字母运算一样。特别约定

$$(x+\mathrm{i}y)(x'+\mathrm{i}y')=(xx'-yy')+\mathrm{i}(xy'+x'y)。$$

由此显而易见，一切乘法法则皆成立，只有单调性法则不加考虑。

(5) 除法定义为乘法之逆，特别是不难证实

$$\frac{1}{x+\mathrm{i}y}=\frac{x}{x^2+y^2}-\mathrm{i}\frac{y}{x^2+y^2}$$

除了对于 $x=y=0$ 以外，这个数永远存在，即像在实数域内一样，不允许用零做除数。

由此可见，复数运算不会导致矛盾，因为复数运算仅只依赖于实数及已知的实数运算。我们在这里假定这些运算都不包含矛盾。

除了这种纯形式的解释以外，我们当然更愿意对复数及复数运算作出几何解释，或作出更看得见的解释，使我们能用图形看到运算法则相容性基础。常用的几何解释就提供了这样的基础。正如你们都已了解，也正如我们已提到过的那样，几何解释用平面上 xy 坐标系的点(x, y)代表复数 $z=x+iy$。相应的两个数 z 与 a 的和可用熟知的平行四边形法则求出，即用这两个点和原点 O 作平行四边形，而乘积 $z\cdot a$ 则可在线段 Oz 上作一相似于 $aO1$ 的三角形求得，其中 1 是一个点 ($x=1$, $y=0$)（图 4.1）。简言之，加法 $z'=z+a$ 用平面本身的平移来表示。乘法 $z'=z\cdot a$ 用相似变换来表示，即使原点保持固定，作一旋转和延伸。平面上的点代表复数，从点的排列上可以立即看出，这里是什么代替了实数的单调性定律。我希望这些解释已足以使你们清楚地回忆起学过的内容。

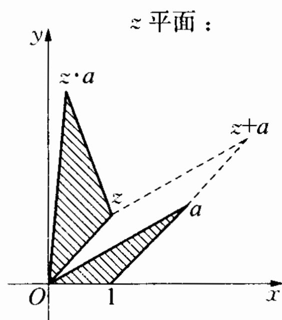


图 4.1

最后，我必须提请你们注意对复数的几何解释在高斯研究中所占的突出地位，因为复数的这个基础正是由于几何解释才得到了充分的强调。而复数的普遍重要性也正是由此而第一次被表现了出来。1831 年，这个研究使高斯进入了这个理论，特别是整复数 $a+ib$（其中 a, b 是实的整数）的理论领域，他对于这种新数发展了通常数论中有关素因数、二次剩余、双二次剩余等定理。数论的这种推广，我们曾结合费马定理的讨论提到过。在他自己写的那篇文章摘要中①，高斯曾就他所谓的虚数的真义表示过意见。对于他来说，复数运算的正确性，是由人们赋予它及其运算的几何解释，才得到证实的。因此，他所采取的立场决不是形式的立场。高斯的

① 见《高斯著作集》第 2 卷。

这些很长的写得很漂亮的文章是非常值得一读的。我在这里再提一句，高斯使用了更清楚的"复数"一词以代替"虚数"，复数这个名词从此以后就被采用了。

§4.2 高阶复数，特别是四元数

每一个严肃研究过复数的人都会想到要问：复数中带有一个虚数单位 i，我们能不能建立带有一个以上的"单位"的更高阶的复数？我们能不能对它们进行合于逻辑的运算？大约在 1840 年左右，H·格拉斯曼在斯德丁和 W·R·哈密顿在都柏林各自独立地取得了这方面的成果。后面我们要详细探讨哈密顿的发明——四元数的演算。现在让我们来看一般问题。

我们可以把一般的复数 $x+\mathrm{i}y$ 看做是线性组合

$$x\cdot 1+y\cdot \mathrm{i}。$$

它是由两个不同的单位数 1 和 i 通过实参数 x 和 y 组成的。同样地，现在让我们想象有任意 n 个不同的单位数 $e_1, e_2, \cdots, e_n$ 组成，并把它们的一切组合 $x=x_1e_1+x_2e_2+\cdots+x_ne_n$ 称为带有 n 个任意实数 $x_1, x_2, \cdots, x_n$ 的高阶复数系。如果给出两个这样的数，一个是上面定义的 x，还有

$$y=y_1e_1+y_2e_2+\cdots+y_ne_n,$$

那么几乎是显而易见地，当且仅当各个单位数的系数（即该数的诸"分量"）一一相等时，

$$x_1=y_1, x_2=y_2, \cdots, x_n=y_n,$$

我们才称这两个数相等。其加减运算简单地归纳为其中系数的加减运算

$$x\pm y=(x_1\pm y_1)e_1+(x_2\pm y_2)e_2+\cdots+(x_n\pm y_n)e_n,$$

这也同样地明显。

就乘法的情况来说，事情就比较困难，也比较有趣了。首先我们从字母相乘的一般规则谈起，即把 x 中的每一个 i 项乘上 y 中的每一个 k 项 $(i, k = 1, 2, \cdots, n)$。由此给出

$$x \cdot y = \sum_{(i,\ k=1,\ 2,\ \cdots,\ n)} x_i y_k e_i e_k 。$$

为了使这个式子成为我们所说的数系中的一个数，必须有一个规则能把乘积 $e_i \cdot e_k$ 表示为此系中的复数，即表示为单位数的线性组合，因而必须有 n^2 个方程如下：

$$e_i e_k = \sum_{(l=1,\ \cdots,\ n)} c_{ikl} \cdot e_l \quad (i, k = 1, \cdots, n)。$$

这样就可以说，数

$$x \cdot y = \sum_{(i=l,\ \cdots,\ n)} \left\{ \sum_{(i,\ k=1,\ \cdots,\ n)} x_i y_k c_{ikl} \right\} e_l$$

将永远属于我们所说的复数系统。各个特定的复数系统，都是由确定这个乘法规则的方法，即根据系数 c_{ikl} 的表来刻画。

如果把除法定义为乘法之逆运算，那么在这样的一般定义之下，就会发现除法并不一定是唯一可能的，即使除数不等于零。因为确定 $x \cdot y = z$ 中的 y，要求对 n 个未知的 $y_1, \cdots, y_n$ 解 n 个线性方程 $\sum\limits_{i,k} x_i y_k c_{ikl} = z_l$，如果行列式正好是等于零，那么结果或者无解，或者有无限多个解，而且所有的 z_l 可能是 0，就是说，即使并非所有的 x_i 或所有的 y_k 等于零，即两个数的乘积可以等于 0 而因子却都不等于 0。只有巧妙地特别选择系数 c_{ikl}，才能使之与通常的数的行为一致。确实，仔细研究一下就可以了解，当 $n > 2$ 时，为了达到这一点，我们必须牺牲某一个运算规则。在这种情况下，我们选择一个看来不那么重要的规则，作为不能被满足的规则。

在作了这些一般性的说明以后，我们以四元数作为一个例子来

详细加以讨论。因为四元数在物理和数学中皆有应用，所以它是最重要的高阶复数系统。正如术语名称所表示的，它是四项数（$n=4$）；它包括三维向量作为一个子类，三维向量今天已得到普遍了解，有时在中学里也加以讨论。

在我们用来构造四元数的 4 个单位数中，我们将选实数单位 1 作为第一个（像普通复数的情况一样）。其他 3 个单位像哈密顿一样通常记作 i，j，k，所以四元数一般记作

$$p = d + \mathrm{i}a + \mathrm{j}b + \mathrm{k}c,$$

其中 a，b，c，d 是作为四元数系数的实参数。它的第一个分量，即乘以 1 的分量，对应于一般复数的实数部分，我们称之为四元数的“标量部分”，把其他 3 项的组合 $a\mathrm{i}+b\mathrm{j}+c\mathrm{k}$ 称为“向量部分”。

四元数的加法可以从前述一般原则推出。我将对此提出一个显而易见的几何解释，这个解释来自你们已熟悉的对向量的解释。想象有一线段对应于 p 的向量部分，它在坐标轴上的投影为 a，b，c，再想象其上有一个等于标量部分的重量。于是 p 和 $p'=d'+\mathrm{i}a'+\mathrm{j}b'+\mathrm{k}c'$ 相加，首先根据熟知的向量相加的平行四边形法则，作出两个向量的合向量（图 4.2），再在其上加一个重量等于原来两重量之和，这样做实际上就表示了四元数

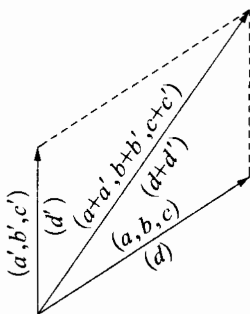


图 4.2

$$p+p'=(d+d')+\mathrm{i}(a+a')+\mathrm{j}(b+b')+\mathrm{k}(c+c')。\quad (1)$$

当转到乘法的时候就第一次接触到四元数的特殊性质了。正如一般情况下所见到的那样，这些性质必包含在有关单位数乘积的约定中。首先我要提出 16 个四元数，按哈密顿规定，它们是那 4 个单位两两相乘的乘积。如符号所示，我们把第一个单位数 1 当作实数 1 来运算，所以

$$1^2=1,\ \mathrm{i}\cdot 1=1\cdot\mathrm{i}=\mathrm{i},$$
$$\mathrm{j}\cdot 1=1\cdot\mathrm{j}=\mathrm{j},\ \mathrm{k}\cdot 1=1\cdot\mathrm{k}=\mathrm{k}。\tag{2a}$$

然后我们规定,其他单位数的平方均为−1,这就是本质上新的东西

$$\mathrm{i}^2=\mathrm{j}^2=\mathrm{k}^2=-1。\tag{2b}$$

对于其二元乘积则规定

$$\mathrm{jk}=+\mathrm{i},\ \mathrm{ki}=+\mathrm{j},\ \mathrm{ij}=+\mathrm{k},\tag{2c}$$

而对于因子的换位则规定

$$\mathrm{kj}=-\mathrm{i},\ \mathrm{ik}=-\mathrm{j},\ \mathrm{ji}=-\mathrm{k}。\tag{2d}$$

这里我们看到一个令人惊讶的情况：不服从乘法的交换律。这是四元数的不便之点,但我们必须接受它,才能保持除法的唯一性,并使仅当因子之一为零时乘积才为零的定理成立。我们立刻可以看到,不仅这个定理成立,而且一切加法和乘法法则除了一个例外,都仍然成立。换一句话说,这些简单的约定是非常方便的权宜办法。

我们先作出两个一般四元数

$$p=d+\mathrm{i}a+\mathrm{j}b+\mathrm{k}c\text{ 及 }q=w+\mathrm{i}x+\mathrm{j}y+\mathrm{k}z$$

的乘积。让我们从下列方程出发

$$q'=p\cdot q=(d+\mathrm{i}a+\mathrm{j}b+\mathrm{k}c)\cdot(w+\mathrm{i}x+\mathrm{j}y+\mathrm{k}z),$$

并且逐项相乘。乘的时候必须注意单位 i, j, k 的顺序。我们必须遵循对系数 a, b, c, d 的乘积的交换律,以及对系数和单位 1 乘积的交换律。必须根据乘法表代入单位的乘积,然后再合并有相同单位的同类项。于是我们有

$$\left.\begin{aligned}q'=p\cdot q=w'+\mathrm{i}x'+\mathrm{j}y'+\mathrm{k}z'=&(dw-ax-by-cz)\\&+\mathrm{i}(aw+dx\underline{+bz-cy})\\&+\mathrm{j}(bw+dy\underline{+cx-az})\\&+\mathrm{k}(cw+dz\underline{+ay-bx})。\end{aligned}\right\}\tag{3}$$

于是,四元数乘积的分量,就简单地确定为两个因式分量的双线性组合。如果倒置因数的顺序,那么下方画杠杠的 6 项就会改变符号,因此 $q \cdot p$ 一般不同于 $p \cdot q$,而且区别与单位的乘积不同,也不止是符号的改变。

虽然交换律对乘法不成立,但分配律和结合律仍然成立。因为如果形式上相乘而不代入单位数的乘积,一方面定 $p(q+q_1)$;另一方面定 $pq+pq_1$,那么必然得到相同的结果,以后用乘法表代入乘积也不会有变化。再进一步说,如果结合律对于单位数相乘成立的话,那么对于一般四元数也成立。

但是正如下例所表明的,从乘法表立刻得出

$$(\mathrm{ij})\mathrm{k} = \mathrm{i}(\mathrm{jk})。$$

实际上,我们有

$$(\mathrm{ij})\mathrm{k} = \mathrm{k} \cdot \mathrm{k} = -1,$$

及

$$\mathrm{i}(\mathrm{jk}) = \mathrm{i} \cdot \mathrm{i} = -1。$$

下面来讨论除法。我们只需表明,对于每一个四元数 $p = d + \mathrm{i}a + \mathrm{j}b + \mathrm{k}c$,存在着另一个确定的四元数 q,使

$$p \cdot q = 1。$$

我们可以把 q 记作$\frac{1}{p}$。正如我们下面所表明的,一般的除法可以容易地划归为这个特例。为了确定 q,让我们在方程(3)中设

$$q' = 1 = 1 + 0 \cdot \mathrm{i} + 0 \cdot \mathrm{j} + 0 \cdot \mathrm{k},$$

使系数分别相等,对于 q 中的未知系数 x, y, z, w,求得下列 4 个方程

$$\begin{aligned} dw - ax - by - cz &= 1, \\ aw + dx - cy + bz &= 0, \\ bw + cx + dy - az &= 0, \\ cw - bx + ay + dz &= 0。\end{aligned}$$

正如大家所熟知的，这样一组方程的可解性取决于它的行列式，而这个行列式是斜对称行列式，其中主对角线的一切元素相同，而关于对角线成对称位置的各对元素，绝对值都相等而符号相反。根据行列式理论，这样的行列式是很容易计算的。我们得

$$\begin{vmatrix} d & -a & -b & -c \\ a & d & -c & b \\ b & c & d & -a \\ c & -b & a & d \end{vmatrix} = (a^2+b^2+c^2+d^2)^2。$$

通过直接计算，这个结果不难证实。哈密顿的约定真正漂亮的地方，就是这个结果，即这个行列式是 p 的 4 个系数平方和之乘方。因而由此得出，除非 $a=b=c=d=0$，这个行列式永远不同于 0。除了这个不言而喻的例外 $(p=0)$，上述方程组都有唯一解，所以非 0 四元数的倒数是唯一确定的。

量

$$T=\sqrt{a^2+b^2+c^2+d^2}$$

在这个理论中起着重要作用，称为 p 的张量。不难证明，上述方程组的唯一解是

$$x=-\frac{a}{T^2},\ y=-\frac{b}{T^2},\ z=-\frac{c}{T^2},\ w=\frac{d}{T^2},$$

从而最后结果为

$$\frac{1}{p}=\frac{1}{d+\mathrm{i}a+\mathrm{j}b+\mathrm{k}c}=\frac{d-\mathrm{i}a-\mathrm{j}b-\mathrm{k}c}{a^2+b^2+c^2+d^2}。$$

如果我们像在普通复数中一样引入 p 的共轭值：

$$\bar{p}=d-\mathrm{i}a-\mathrm{j}b-\mathrm{k}c\ ①,$$

---

① 英译本作 $P=d-\mathrm{i}a-\mathrm{j}b-\mathrm{k}c$，是不对的，$P$ 应该为 $\bar{p}$。——中译者

那么我们可以把上面的公式记作

$$\frac{1}{p}=\frac{\overline{p}}{T^{2}},$$

或

$$p\cdot\overline{p}=T^{2}=a^{2}+b^{2}+c^{2}+d^{2}。$$

这些公式是普通复数的某些性质的直接推广。既然 p 也是共轭于$\overline{p}$的数,则也可得出:

$$\overline{p}\cdot p=T^{2}。$$

所以在这一个特例中交换律成立。

现在除法的一般问题也就可以解决了。因为方程

$$p\cdot q=q'$$

如用$\frac{1}{p}$来乘,得

$$q=\frac{1}{p}\cdot q'=\frac{\overline{p}}{T^{2}}\cdot q',$$

而改变因子的顺序所得到的方程

$$q\cdot p=q'$$

的解,则是

$$q=q'\cdot\frac{1}{p}=q'\cdot\frac{\overline{p}}{T^{2}}。$$

这个解一般不同于前一个解。

现在我们必须问:对这些四元数的运算及其运算法有没有几何解释,使它们出现得很自然?为了作出回答,我们从两个因数都可以简化为简单向量的特例出发,即从标量部分 w, d 为零的特例出发。因而用于乘法的公式(3)变成

$$
\begin{aligned}
q' = p \cdot q &= (\mathrm{i}a + \mathrm{j}b + \mathrm{k}c)(\mathrm{i}x + \mathrm{j}y + \mathrm{k}z) \\
&= -(ax + by + cz) + \mathrm{i}(bz - cy) + \mathrm{j}(cx - az) \\
&\quad + \mathrm{k}(ay - bx),
\end{aligned}
$$

即当两个四元数各简化为一个向量时，它们的乘积由一个标量部分和一个向量部分构成。这两个部分就不难同现在使用的各种向量乘法联系起来。向量运算的概念比四元数运算用的广得多，应追溯到格拉斯曼。但"向量"这个词是英语中来的。与上述乘积相关联的有两种向量的乘法：其一，现在多半称作内积（或标量积）$ax + by + cz$（即除符号以外为上述四元数乘积的标量部分）；其二称为外积（或向量积）$\mathrm{i}(bz - cy) + \mathrm{j}(cx - az) + \mathrm{k}(ay - bx)$（即四元数乘积的向量部分）。下面我们分别对每个部分作出几何解释。

让我们用从原点 O 出发的两条线段表示两个向量 (a, b, c) 及 (x, y, z)（图 4.3）。它们相应地以 (a, b, c) 及 (x, y, z) 作为终点，具有长度 $l = \sqrt{a^2 + b^2 + c^2}$ 及 $l' = \sqrt{x^2 + y^2 + z^2}$。如果 φ 是这两线段之间的角，那么众所周知，毋需我在这里再多讲，由解析几何公式，其内积为

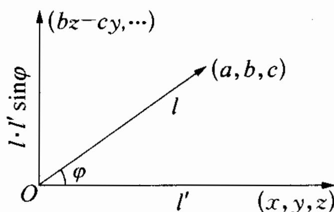


图 4.3

$$
ax + by + cz = l \cdot l' \cdot \cos\varphi;
$$

其外积本身为一向量，这个向量很容易看出垂直于 l 及 l' 所在的平面，具有长度 $l \cdot l' \cdot \sin\varphi$。

现在最重要的是确定外积向量的指向，即向着 l 及 l' 所在平面上的哪一侧画出这个向量。根据所选择的坐标系，这个指向是不同的。正如你们所知道的，可以选择不是全等的两个直角坐标系，即将 y 轴及 z 轴固定，倒转 x 轴的指向才能互相叠合的

直角坐标系[①]。这样，两个坐标系就互为对称，如图 4.4 中的右手和左手。关于两者之间的区别，可记住以下规则：在一个坐标系中，x，y，z 轴相当于右手向前伸着的拇指、食指、中指；在另一个坐标系中，相当于左手向前伸着的拇指、食指、中指。在文献中，这两个坐标系搞得很乱，因为各个国家、各个行业、各个作者，甚至同一个作者习惯都不同。现在让我们来看一个最简单的例子，其中 $p=\mathrm{i}$，$q=\mathrm{j}$，令其为 x 轴和 y 轴上画出的单位长度。既然 $\mathrm{i}\cdot\mathrm{j}=\mathrm{k}$，则向量的外积为 z 轴上画出的单位长度(图 4.5)。现在可以把 i 和 j 连续地转换为两个任意向量 p 和 q，于是 k 不经过零而连续地转换为 $p\cdot q$ 的向量部分。因而第一个向量、第二个向量及向量乘积之间的相对关系，相当于右手(图 4.4)或左手(图 4.4)坐标系中的 x 轴、y 轴、z 轴[②]，因所选择的坐标系而异(德国的习惯如图 4.5 所示)。

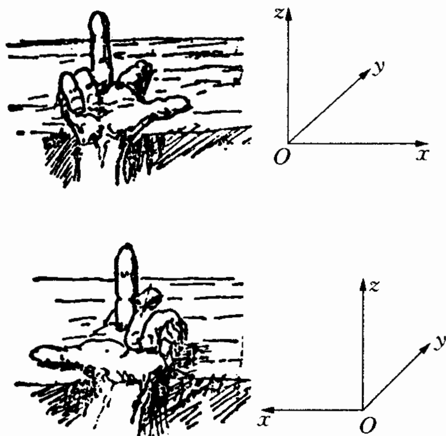


图 4.4

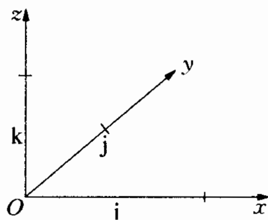


图 4.5

最后，我想对争论很多的向量分析记号问题补充几句话。用于向量运算的符号很多很多，各不相同，至今不能提出一个能被普遍接

① 英译文作倒转 x 轴也不能叠合，是不对的。——审订者

② 用图 4.4 说明问题似更为恰当，英译本用图 4.5(右手)、图 4.3(左手)。——编辑

受的记号。为此,1903 年在卡塞尔举行的自然科学家大会上,建立了一个委员会。但是委员会成员之间甚至不能互相完全了解。尽管他们的用心都是好的,但是各个成员都不肯完全俯就他人的意见,以至最后又搞出了三四个新的记号!我在这些事情上的经验使我倾向于认为,只有在重大的物质利益的要求下才能达成一致的协议。例如电工学上的伏特、安培、欧姆之类统一的度量单位,是到了 1881 年才被一致采纳的,后来又以公共立法的形式固定下来。究其原因,就是由于有了那种压力,因为工业上迫切需要那种统一,作为一切计算的基础。但是在向量运算的背后却不存在那种强烈的物质利益,至少目前还是这样,因而我们只好让数学家各抱一个他认为最方便的记号,如果他主观认为他用的符号是唯一正确的,那么好坏也由它去吧。

§4.3　四元数的乘法——旋转和伸展

在考虑一般四元数乘法的几何意义之前,让我们先考虑以下的问题。我们来考虑两个四元数 p 和 q 的乘积 $q'=p\cdot q$,并以其共轭 $\overline{p}$ 和 $\overline{q}$ 来替代 p 和 q,即改变符号 a, b, c, x, y, z。因而如上一节方程(3)所给出的,乘积的标量部分保持不变,只有 i, j, k 的那些下方未画杠杠的系数改变符号。但是如果我们颠倒因数 $\overline{p}$ 和 $\overline{q}$ 的顺序,那么 i, j, k 的下方画了杠杠的系数就会改变符号。因而乘积 $\overline{q'}=\overline{q}\cdot\overline{p}$ 正好是原乘积 q' 的共轭,得

$$q'=p\cdot q,\ \overline{q'}=\overline{q}\cdot\overline{p},$$

其中 $\overline{q'}$ 是 q' 的共轭。把这两个方程相乘,得

$$q'\cdot\overline{q'}=p\cdot q\cdot\overline{q}\cdot\overline{p}。$$

在这个方程中,因数的顺序是重要的,因为交换律不成立。不过可以应用结合律,记作

$$q' \cdot \overline{q'} = q \cdot (q \cdot \overline{q}) \cdot \overline{p}。$$

因按上一节方程得

$$q \cdot \overline{q} = x^2 + y^2 + z^2 + w^2,$$

故可记作

$$w'^2 + x'^2 + y'^2 + z'^2 = p(w^2 + x^2 + y^2 + z^2)\,\overline{p}。$$

右边中间的因数是一个标量，而对于四元数乘以标量交换律成立，由于 $M \cdot p = Md + \mathrm{i}(Ma) + \mathrm{j}(Mb) + \mathrm{k}(Mc) = pM$，因而有

$$w'^2 + x'^2 + y'^2 + z'^2 = p\,\overline{p}(w^2 + x^2 + y^2 + z^2)。$$

又因 $p \cdot \overline{p}$是 p 的张量的平方，我们有

$$\begin{aligned} &w'^2 + x'^2 + y'^2 + z'^2 \\ &= (d^2 + a^2 + b^2 + c^2)(w^2 + x^2 + y^2 + z^2), \end{aligned} \qquad (\mathrm{I})$$

即两个四元数乘积的张量等于其因数张量的乘积[①]。这个公式也可以从上一节给出的乘积公式中取 w', x', y', z'的值通过直接运算求得。

现在我们把一个四元数表示为连接四维空间的原点和点(x, y, z, w)的一个线段，方式与三维空间的向量表示完全相同。我当大学生的时候利用四维空间是感到惴惴不安的，因为当时习惯如此，现在不再必要了。你们大家都完全清楚其中没有玄学的意义，因为高维空间不过是一种方便的数学表示，使我们能利用与真实空间表示的相类似的术语罢了。如果我们把 p 看做是一个常数，即把 a, b, c, d 看做是常数，那么四元数方程

$$q' = p \cdot q$$

就表示四维空间点(x, y, z, w)到另一点(x', y', z', w')的某一线

① 这个公式的一切要点，见拉格朗日(L. Lagrange)的著作。

性变换，因为这个方程使每一个四维向量 q 线性地变为另一个向量 q'。这个变换的显式方程，即作为 x, y, z, w 线性函数的 x', y', z', w' 的表示式，可以从与前一节乘积公式(3)比较系数而求得。张量方程(Ⅰ)表明，任一点与原点的距离 $\sqrt{x^2+y^2+z^2+w^2}$，对于空间上的一切点，都是乘以相同的常数因子 $T=\sqrt{a^2+b^2+c^2+d^2}$。最后，根据前一节所述的行列式的计算结果，这种线性变换的行列式显然是正确的。

三维空间的解析几何说明，如果坐标 x, y, z 的线性变换是正交的(即如果它把式 $x^2+y^2+z^2$ 仍变为本身)，并且如果变换的行列式是正的，那么变换所表示的是围绕原点的旋转。反过来，任何旋转都可以这种方式求得。但是如果线性变换把 $x^2+y^2+z^2$ 变成乘以常数因子 T^2 的 x', y', z' 的类似的式子，如果行列式是正的，那么变换所表示的就不但是围绕原点的旋转，而且附加一个关于原点按比例 T 所作的伸展，或简单地称为旋转与伸展。

刚才提到的三维空间的情况，可以推广到四维空间。我们这种四维空间的变换，在完全相同的意义下表示绕原点的旋转及伸展。不过很容易看出，在这种情况下所求得的，不是最一般的绕原点的旋转和伸展，因为这个变换只包含 4 个任意常数，即 p 的系数 a, b, c, d。但我们马上要说明，围绕四维空间 R_4 原点的最一般的旋转和伸展，包含 7 个任意常数。的确，为了使一般线性变换成为旋转和伸展，必须有

$$x'^2+y'^2+z'^2+w'^2=T^2(x^2+y^2+z^2+w^2)。$$

如果我们用 x, y, z, w 的线性整函数来代替 x', y', z', w'，那么就得到一个具有 4 个变量的二次型，其中含有 $\frac{4\cdot 5}{2}=10$ 项。使系数相等的结果，得 10 个方程式。由于 T 仍然是任意的，所以 16 个系数中只有 9 个方程，剩下 7 个任意常数。

值得指出的是：尽管如此，最一般的旋转和伸展仍可以通过四

元数的乘法求得。令 $\pi=\delta+\mathrm{i}\alpha+\mathrm{j}\beta+\mathrm{k}\gamma$ 为另一个常四元数。这样就可以像刚才一样证明,与前一个变换仅只顺序不同的变换 $q'=q\cdot\pi$,可表示 R_4 的一个旋转和伸展。因而复合变换

$$\begin{aligned} q' &= p\cdot q\cdot \pi \qquad\qquad (\mathrm{II})\\ &= (d+\mathrm{i}a+\mathrm{j}b+\mathrm{k}c)\cdot q\cdot(\delta+\mathrm{i}\alpha+\mathrm{j}\beta+\mathrm{k}\gamma)\end{aligned}$$

也表示一旋转和伸展。这一变换仅包含 7 个(不是 8 个)任意常数。因为如果把 a, b, c, d 乘以任何实数,把 α, β, γ, δ 除以同样的数,那么这个变换仍然不变。因而这个组合变换很可能(plausible)表示着四维空间的最一般旋转和伸展。凯莱(Cayley)证明了这个漂亮的结果确实是对的。我只想提一下这个历史事实,以免讲得太琐碎。这个公式出现在 1854 年凯莱所写的文章《二阶曲面到自身的等画自变换》以及他的其他某些文章中。

凯莱的这个公式有一个很大的好处,就是能使我们立刻理解两个旋转和伸展的组合。例如,若根据下列方程给出第二个旋转和伸展

$$q''=w''+\mathrm{i}x''+\mathrm{j}y''+\mathrm{k}z''=p'\cdot q'\cdot\pi',$$

其中 p'和 π'是已知的新的四元数,则根据(Ⅱ)得

$$q''=p'\cdot(p\cdot q\cdot\pi)\cdot\pi'。$$

由此据结合律

$$q''=(p'\cdot p)\cdot q\cdot(\pi\cdot\pi'),$$

或

$$q''=\gamma\cdot q\cdot\varrho,$$

其中 $\gamma=p'\cdot p$ 和 $\varrho=\pi\cdot\pi'$ 是新给定的四元数。因而得一旋转和伸展式,这个式子正像过去的形式一样把 q 转换为 q''。我们看到,四元数乘积中 q 的前后乘数分别为单独的变换中 q 的对应乘数的乘积,

而诸因数的顺序必须如公式所示。

这种四维表示可能看来不够令人满意，可能会希望得出一个可以用通常的三维空间来表示的东西。所以，我们将证明可以利用刚才给出的公式的一个特款，为类似的三维运算求得类似的公式。确实，四元数乘法对于一般物理和力学的重要性，正是建立在这些公式的基础上的。我说"对于一般物理和力学"，是因为我不想在这里阐述前述公式可以不加修正地应用于这些学科的推广。那些推广比你们所能想象的更直接。与相对论原则相联系的电动力学的新发展，在本质上不外乎是旋转和伸展在四维空间中的合乎逻辑的应用结果。这些概念最近已由闵可夫斯基提出和推广①。

我们还是保持在三维空间的范围内。在三维空间中，旋转和伸展把点(x, y, z)变换为点(x', y', z')，致使

$$x'^2 + y'^2 + z'^2 = M^2(x^2 + y^2 + z^2),$$

其中M表示每一长度伸展的比。由于从(x, y, z)到(x', y', z')的一般线性变换包含 9 个系数，上述方程的左边在代入x', y', z'值后成为x, y, z的二次型，其中包含 6 项，因此将上述方程中的系数加以比较，得出 6 个方程。如果假设M值为任意的，那么就减少到 5 个方程。而在线性变换中原有的 9 个系数，因为要服从于这 5 个条件，所以减少到只有 4 个是任意常数(对照前几页所述)。如果这样的线性变换有一个正行列式，那么如前几页所示，它所表示的是绕原点的空间旋转，同时以$\frac{1}{M}$之比伸展。如果行列式是负的，那么变换所表示的是旋转和伸展与一个反射相复合，例如，由方程$x = -x'$, $y = -y'$, $z = -z'$所定义的就是一个反射。此外可以证明，变换的

① 自此处写成后，上述特殊的相对论领域已发表了大量的文章。这里请允许我提一下我的一次讲演："Über die geometrische Grundlagen der Lorentzgruppe"，刊于 *Jahresbericht der deustchen mathematiker — Vereingung*，第 19 卷，第 299 页，1910 年，重印于克莱因 *Gesammelte mathematische Abhandlungen*，第 1 卷，第 533 页。

行列式必取$\pm M'$两值之一。

为了用四元数来表示这些关系，让我们首先把四元变数 q 和 q' 退化为其向量部分

$$q' = \mathrm{i}x' + \mathrm{j}y' + \mathrm{k}z',\ q = \mathrm{i}x + \mathrm{j}y + \mathrm{k}z。$$

我们把这些看做是连接原点和变换前后点的位置的三维向量。我们现在要证明三维空间的一般旋转和伸展仍由公式(Ⅱ)给出，但其中 p 和 π 具有共轭值，即 $q' = p \cdot q \cdot \bar{p}$，或记为展开形式

$$\begin{aligned} &\mathrm{i}x' + \mathrm{j}y' + \mathrm{k}z' \\ &= (d + \mathrm{i}a + \mathrm{j}b + \mathrm{k}c)(\mathrm{i}x + \mathrm{j}y + \mathrm{k}z)(d - \mathrm{i}a - \mathrm{j}b - \mathrm{k}c)。 \end{aligned} \tag{1}$$

为了证明这一点，我们必须首先说明，右边乘积的标量部分等于零，即 q'确为向量。要做到这一点，先按四元数的乘法法则将 p, q 相乘，得

$$\begin{aligned} q' = [&-ax - by - cz + \mathrm{i}(dx + bz - cy) + \mathrm{j}(dy + cx - az) \\ &+ \mathrm{k}(dz + ay - bx)] \cdot (d - \mathrm{i}a - \mathrm{j}b - \mathrm{k}c), \end{aligned}$$

再乘另一四元数，实际计算得 q'的标量部分为零，而对向量部分的各分量得出表达式

$$\begin{cases} x' = (d^2 + a^2 - b^2 - c^2)x + 2(ab - cd)y + 2(ac + bd)z, \\ y' = 2(ba + cd)x + (d^2 + b^2 - c^2 - a^2)y + 2(bc + ad)z, \\ z' = 2(ca - bd)x + 2(cb + ad)y + (d^2 + c^2 - a^2 - b^2)z。 \end{cases} \tag{2}$$

这些公式实际上代表着一个旋转和伸展，如果把(1)写作张量方程，这就很明显了。这个张量方程根据公式(Ⅰ)为

$$\begin{aligned} &x'^2 + y'^2 + z'^2 \\ &= (d^2 + a^2 + b^2 + c^2)(x^2 + y^2 + z^2)(d^2 + a^2 + b^2 + c^2), \end{aligned}$$

或

$$x'^2 + y'^2 + z'^2 = T^4(x^2 + y^2 + z^2),$$

其中 $T = \sqrt{d^2 + a^2 + b^2 + c^2}$ 表示 p 的张量。因此，如果行列式为

正，那么这个变换正是一个旋转和伸展（见前），否则就是与反射相复合的变换。无论在哪一种情况下，伸展比都是 $M=T^2$。如前面所指出的，在两个值 $\pm M^3=\pm T^6$ 中，行列式必取其一。如果我们考虑对应于同一张量值 T（显然必须不等于零）的参数 a, b, c, d 的一切可能值，考虑与这样的参数相对应的变换，那么就看到，如果对于 a, b, c, d 的任何单独一组值，行列式之值为 $+T^6$ 的话，则此行列式恒取值 T^6，因为这个行列式是 a, b, c, d 的连续函数，因而它不可能不取中间值而突然由 $+T^6$ 变成 $-T^6$。使行列式为正的一组值是 $a=b=c=0, d=T$，因为据(2)，a, b, c, d 这些值的行列式值为

$$\begin{vmatrix} d^2 & 0 & 0 \\ 0 & d^2 & 0 \\ 0 & 0 & d^2 \end{vmatrix} = d^6 = +T^6。$$

由此可见符号恒为正，实际上(1)式总表示一个旋转和伸展。所以不难写出把反射同旋转和伸展结合起来的变换，因为只需要把前一变换同反射 $x'=-x, y'=-y, z'=-z$ 结合起来就行了，它等价于四元数方程 $\overline{q'}=p\cdot q\cdot \overline{p}$。

我们现在要证明，反过来，每一个旋转和伸展可以记作形式(1)，或记作等价的形式(2)。首先，如前所述，这个公式包含了一般情况所必须具有的 4 个任意常数。我们可以通过对这 4 个常数的适当选择，以实际求得任何所要求的伸展比值 $M=T^2$，任何所要求的旋转轴的位置，以及任何所要求的旋转角。令 ξ, η, ζ 表示旋转轴的方向余弦，令 ω 表示旋转角。当然，我们有众所周知的关系

$$\xi^2+\eta^2+\zeta^2=1。 \tag{3}$$

现在要证明 a, b, c, d 系由方程(4)给出

$$\begin{cases} d=T\cdot\cos\dfrac{\omega}{2}, \\ a=T\cdot\xi\cdot\sin\dfrac{\omega}{2},\ b=T\cdot\eta\cdot\sin\dfrac{\omega}{2},\ c=T\cdot\zeta\cdot\sin\dfrac{\omega}{2}。 \end{cases} \tag{4}$$

根据(3)式,方程(4)显然满足条件

$$d^2+a^2+b^2+c^2=T^2。$$

在这些关系得到证明以后,显然可以对 T, ξ, η, ζ, ω 的任何给定值求得 a, b, c, d 的正确值。

为了证明关系(4),让我们首先指出,如果给定 a, b, c, d,那么就确定了量 ω, ξ, η, ζ,使(3)式得到满足。因为 T 是四元数 $p=d+\mathrm{i}a+\mathrm{j}b+\mathrm{k}c$ 的张量,所以对方程(4)平方相加,得

$$1=\cos^2\frac{\omega}{2}+\sin^2\frac{\omega}{2}(\xi^2+\eta^2+\zeta^2)。$$

因而(3)式成立。由此得出,ξ, η, ζ 完全由关系(4′)确定

$$a:b:c=\xi:\eta:\zeta \tag{4'}$$

这是直接从关系(4)来的。这些方程表明点(a, b, c)处于变换的旋转轴上。这个事实是容易验证的,因为设(2)中 $x=a$, $y=b$, $z=c$,即得

$$x'=(d^2+a^2+b^2+c^2)a=T^2\cdot a,$$
$$y'=(d^2+a^2+b^2+c^2)b=T^2\cdot b,$$
$$z'=(d^2+a^2+b^2+c^2)c=T^2\cdot c,$$

即点(a, b, c)仍然在通过原点的同一条射线上,说明它是旋转轴上的一点。剩下来的只是要证明由方程(4)所定义的角真正是旋转角。这个证明需要很长的讨论,不过我可以避免它,只消指出,当 $T=1$ 时的变换(2)就是欧拉对于绕一轴旋转 ω 角所给出的变换,此轴的方向余弦为 ξ, η, ζ。

最后,如果把方程(4)给出的值代入方程(1),则对于绕方向余弦为 ξ, η, ζ 的轴所作的转动角 ω 的旋转,再复合一个比为 T^2 的伸展,得出非常简便的用四元数表示的方程,即为

$$\begin{cases} \mathrm{i}x' + \mathrm{j}y' + \mathrm{k}z' = T^2\left\{\cos\frac{\omega}{2} + \sin\frac{\omega}{2}(\mathrm{i}\xi + \mathrm{j}\eta + \mathrm{k}\zeta)\right\} \\ \quad \cdot \{\mathrm{i}x + \mathrm{j}y + \mathrm{k}z\} \cdot \left\{\cos\frac{\omega}{2} - \sin\frac{\omega}{2}(\mathrm{i}\xi + \mathrm{j}\eta + \mathrm{k}\zeta)\right\} \end{cases} \text{。} \tag{5}$$

这个公式是欧拉旋转公式的一个很容易记忆的形式：向量 $\mathrm{i}x + \mathrm{j}y + \mathrm{k}z$ 前后的乘数相应地为两个共轭的四元数，其张量为单位张量（称为 versor，意为“旋转量”，而与张量相对照，后者表示“伸张的量”），然后将整个结果乘以伸展比的标量因子。

下面将要表明，把这些公式进一步特殊化到二维空间时，它们就成为通过两个复数相乘来表示 xy 平面上的旋转和伸展的著名公式。为此目的，让我们选择旋转轴为 x 轴（$\xi = \eta = 0$，$\zeta = 1$）。因而对于 $z = z' = 0$，公式(5)可以记作

$$\mathrm{i}x' + \mathrm{j}y' = T^2\left(\cos\frac{\omega}{2} + \mathrm{k}\sin\frac{\omega}{2}\right)(\mathrm{i}x + \mathrm{j}y)\left(\cos\frac{\omega}{2} - \mathrm{k}\sin\frac{\omega}{2}\right),$$

或者，充分注意单位之间相乘的法则，记作

$$\begin{aligned} &\mathrm{i}x' + \mathrm{j}y' \\ &= T^2\left\{\cos\frac{\omega}{2}(\mathrm{i}x + \mathrm{j}y) + \sin\frac{\omega}{2}(\mathrm{j}x - \mathrm{i}y)\right\}\left\{\cos\frac{\omega}{2} - \mathrm{k}\sin\frac{\omega}{2}\right\} \\ &= T^2\left\{\cos^2\frac{\omega}{2}(\mathrm{i}x + \mathrm{j}y) + 2\sin\frac{\omega}{2}\cos\frac{\omega}{2}(\mathrm{j}x - \mathrm{i}y)\right. \\ &\quad \left. - \sin^2\frac{\omega}{2}(\mathrm{i}x + \mathrm{j}y)\right\} \\ &= T^2\{(\mathrm{i}x + \mathrm{j}y)\cos\omega + (\mathrm{j}x - \mathrm{i}y)\sin\omega\} \\ &= T^2(\cos\omega + \mathrm{k}\sin\omega)(\mathrm{i}x + \mathrm{j}y)\text{。} \end{aligned}$$

如果两边从右侧乘以因子($-\mathrm{i}$)，即得

$$x' + \mathrm{k}y' = T^2(\cos\omega + \mathrm{k}\sin\omega)(x + \mathrm{k}y)\text{。}$$

这就是两个普通复数相乘的法则，只不过这里用字母 k 代替平常的 i 来记虚数单位$\sqrt{-1}$，可以解释为转动角 ω 的旋转，附加上比为 T^2 的

一个伸展。

让我们回过头来谈三维空间，并修正公式(1)，使它来表示没有伸展的纯粹的旋转。为此必须以 $x'\cdot T^2$, $y'\cdot T^2$, $z'\cdot T^2$来代替 x', y', z'，即以 $q'\cdot T^2$来代替 q'。注意 $p^{-1}=\dfrac{1}{p}=\dfrac{\bar{p}}{T^2}$，可记纯粹旋转的公式为

$$\mathrm{i}x'+\mathrm{j}y'+\mathrm{k}z'=p\cdot(\mathrm{i}x+\mathrm{j}y+\mathrm{k}z)\cdot p^{-1}。\tag{6}$$

不失一般性可假设 p 是一个四元数，其张量为单位张量，即

$$p=\cos\frac{\omega}{2}+\sin\frac{\omega}{2}(\mathrm{i}\xi+\mathrm{j}\eta+\mathrm{k}\zeta)，其中\ \xi^2+\eta^2+\zeta^2=1。$$

由此可见，如果取 T 为单位张量，那么(6)式可由(5)式得来。凯莱在 1845 年第一次提出的公式，就是如此①。

我们可以用一个特别简单的形式来表示两个旋转的复合，就像上面对四维空间所作的表示一样。给出第二次旋转

$$\mathrm{i}x''+\mathrm{j}y''+\mathrm{k}z''=p'(\mathrm{i}x'+\mathrm{j}y'+\mathrm{k}z')p'^{-1},$$

其中

$$p'=\cos\frac{\omega'}{2}+\sin\frac{\omega'}{2}(\mathrm{i}\xi'+\mathrm{j}\eta'+\mathrm{k}\zeta'),$$

旋转轴的方向余弦为 ξ', η', ζ'，旋转角为 ω'，可把最终的旋转变为以下方程

$$\mathrm{i}x''+\mathrm{j}y''+\mathrm{k}z''=p'\cdot p\cdot(\mathrm{i}x+\mathrm{j}y+\mathrm{k}z)\cdot p^{-1}\cdot p'^{-1},$$

因而对于最终的旋转，由下列方程给出旋转轴的方向余弦 ξ'', η'', ζ''及旋转角 ω''

---

① 《论与四元数有关的若干结果》，见其《数学文集》第 1 卷，第 123 页，1889 年。根据凯莱自己说(第 1 卷，第 586 页)，哈密顿已于此前独立地发现了同一公式。

$$p'' = \cos\frac{\omega''}{2} + \sin\frac{\omega''}{2}(i\xi'' + j\eta'' + k\zeta'') = p' \cdot p,$$

因而对绕原点的两次旋转的复合得出了一个简便的式子。相对而言，表示最终旋转的通常的公式就显得相当复杂。由于任何四元数可以表示为一个实数(其张量)和旋转量的乘积，因此也对四元数的乘法作出了一个简单的几何解释，即把它解释为旋转的复合。因此，四元数乘法的不可交换性，与众所周知的绕原点的两个旋转的次序不可交换性的事实相对应。

如果你们想研究四元数表示和应用的历史发展情况，我建议你们去读凯莱本人写的一个极有价值的关于动力学的报告：《关于动力学中若干特殊问题解的进展报告》(1862 年英国科学发展协会报告，印于《凯莱数学文集》，第 4 卷，自第 552 页起，1891 年，剑桥)。

最后我想对四元数的价值和研究情况谈几点一般的意见，作为结束。为此目的，应该把一般四元数运算和简单的四元数乘法法则加以区别。从前面的讨论中可以充分看出，至少后者可以肯定有很大的应用价值。另一方面，像哈密顿构想的那样一般的四元数运算，涉及四元数的加、乘、除运算，其运算步骤实在已达到了任意多的程度。哈密顿就是这样来研究四元数的代数的，由于他也研究了无限过程，因此他可以说已创立了四元数函数理论。又由于交换律对它不成立，因此这种理论的面貌与通常的复变数理论完全不同。公平地说，哈密顿的这些一般的、影响深远的思想还没有得到充分论证，因为还没有找到四元数和其他数学分支之间的必不可少的关系与相互依赖性，也没有得到很重大的应用。因此一般四元数理论并没有引起普遍的兴趣。

但是在数学中和在其他人间事务中一样，有观点冷静而客观的人，也总有形成令人惋惜的个人偏见的人。所以四元数理论也有热情的支持者和激烈的反对者。支持者主要在英美，他们在 1907 年成立了“四元数研究促进会”，后来留美的日本数学家木村(Kimura)又

把这个组织建成了一个完全国际化的机构。有一段时间,罗伯特·波尔(Rebort Ball)爵士担任了这个机构的主席。他们预料通过四元数的紧张研究有可能使数学得到巨大发展。另一方面,也有人对四元数理论一句话也听不进,甚至将四元数乘法这个很有用的概念也拒之门外。据这类人的观点,四元数的一切运算顶多是对 4 个分量的运算,单位元数和乘法表在他们看来是多此一举。在这两个极端之间也有许多人认为,我们应该仔细区别标量和向量。

§4.4 中学复数教学

现在我要抛下四元数的理论,在这一章的结尾谈谈四元数概念在中学课程中的地位。没有人会想要在中学里讲四元数,但一般的复数如 $x+iy$ 却总是要提出来讨论的。我在这里与其长篇大论地对你们讲应该怎么教,不如利用不同时期出的 3 本书告诉你们这种教法的沿革,这样也许更有趣些。

我首先向你们提出 18 世纪下半叶在哥廷根大学占有领导地位的卡斯特纳(Kästner)写的一本书。那时大学里学的还是 19 世纪 30 年代下放到中学里的那些初等数学内容。相应地,卡斯特纳讲的也是初等数学,来听课的人中有大量非数学专业的学生。他那些讲稿的主要材料来自他写的那本书,书名为 *Mathematische Anfangsgründe*(《数学初步》)[①]。这里使我们感兴趣的是第三部第二节:“Anfangsgründe der Analysis endlicher Grö Ben”(“有限量分析初步”)。该书从第 20 页开始讲解虚量,讲法大致如下:“无论是谁,若要求否定量”(当时称“否定量”而不称“负量”)的偶次方根,就是缘木求鱼,因为没有一个“否定量”会是那样的幂。的确,这是很正确的。但是该书第 34 页上又说:“这样的根称为不可能的根或虚根”,接着不作很多求证,就暗

① 第三版,哥廷根,1794 年。

暗地把虚根当作普通的数一样演算起来，尽管刚刚说过不存在虚根。这等于说，子虚乌有的东西加上一个名称就突然好像变成可用的东西一样。你们可以在这里看到莱布尼兹观点的反光，因为根据莱布尼兹的观点，虚数事实上是很可笑的，但尽管如此，却以某种不可理解的方式导出了有用的结果。

卡斯特纳还是一个笔法生动的人，他创造了不少妙语警句，在数学文献中占有相当的地位。现在只举许多例子中的一个。在上述一书的引言中，他仔细讲了“Algebra” (代数)一词的词源。确实，从词冠“al”来看，此词来自阿拉伯语。据卡斯特纳说，代数学家就是化分数为整数者，即处理有理函数并将其简化为公分母的人等。据说最初“Algebra”指的是接骨医生的手术，堂吉诃德就去找“Algebraist”接过肋骨，于是卡斯特纳便引用了堂吉诃德这样的话。当然我不敢断定塞万提斯是否真的采用过这种说法，或许这仅仅是讽刺文字，也未可知。

我向你们提出的第二本书是近一些的著作，是几十年前柏林教授 M · 奥姆(M. Ohm)的书：*Versuch Eines Vollständig Konsequenten Systems der Mathematick*(《建立一致性数学系统的一个尝试》)[①]。这本书有一段时间曾被广泛采用，其目的与卡斯特纳相同。但奥姆极大接近现代的观点，他明确地谈到了数系的扩张原则。例如，他说正像负数一样，$\sqrt{-1}$必须作为一个新的东西加于实数。而且甚至这本书还缺乏几何说明，因为它发表在柯西(A. Cauchy)的划时代著作之前(1831 年)。

最后，在许许多多现代中学课本中，我提出一本广泛采用的书：*Bardeys Anfgabensammlung*[②]。在这本书里，扩张原理占据显著地

① 9 卷本，柏林，1828 年。第 1 卷：《算术与代数》，第 276 页。

② 并可参阅此书的修订本，W · 里茨曼(Lietzmann)及 P · 左尔克(P. Zülke)修订，莱比锡，杜伯纳出版社。还可参阅 H · 法因(H. Fine)著《代数中的数系》以及《大学代数》，分别由希思(Heath)出版公司及吉恩(Ginn)出版公司出版。

位，从而说明了几何解释。这可以看做是今天中学教学的一般立场，尽管在个别地方还停留在早先的发展水平上。这本书所采取的观点，据我看来，已化成最适合中学的教学处理。如果不想用系统发展的叙述而使学生感到厌烦，当然也不进行抽象的逻辑解释，那么我们就应当把复数解释为所熟悉的数的概念的扩张，避免任何神秘感。首要的是应当使学生立刻对复平面上的几何作图说明形成习惯！

本教程中用于讲解算术问题的第一个主要部分，就以此结束。在对代数和分析进行类似的讨论之前，我想插入一大段关于历史问题的附录，以便对目前的一般教法以及拟予改进之处增加新的认识。

附：关于数学的现代发展及一般结构

让我先指出，在至今为止的数学发展史上，可以清楚地看出两个不同的发展过程，这两个过程时而互换位置，时而并行不悖，时而又终于合流。我心里知道有这种区分，但很难用生动的语言来表达，因为当前没有一种区分是恰到好处的。不过如果我告诉你们，人们将怎样按这两个发展过程去编写分析系统的初等部分，通过这个具体例子，你们就能明白我的意思。

按第一个过程来编，简称进程 A，就会编出这样一个系统，这是时下中学初等数学课本中最普遍的作法，其系统如下：

(1) 开头是方程的形式理论，即用有理整函数进行运算，并且处理用根式可以解出的代数方程。

(2) 系统研究幂运算及其逆运算的概念，就会得出对数(已证明对数在数值运算中非常有用)。

(3) 到此为止，我们是把分析和几何完全隔开来看的，但现在就要引进几何概念，给出第二类超越函数即三角函数的定义，它的进一步的理论构成了一门新的独立课程。

(4) 接着就是所谓代数分析，讲授最简单的函数展开成无穷级

数。还考虑一般二项式、对数函数及其逆函数即指数函数,连同三角函数一并考虑。无穷级数及其运算的一般理论同样放在这里。也就在这里出现初等超越函数之间的令人惊讶的关系,特别是著名的欧拉公式

$$e^{ix} = \cos x + i\sin x。$$

这类关系之显得更令人惊讶,是因为其中出现的函数最初是在完全不相关的领域里定义的。

(5) 继中学的这个结构之后,与之相同的发展就是学习魏尔斯特拉斯复变函数理论,它以幂级数的性质作为起点。

下面就让我们反其道而行之,非常简略地按数学的第二个发展过程来讲,我把这个过程简称为进程 B。它的中心思想是解析几何思想,探索数的观念与空间观念的结合。

(1) 我们从最简单的函数、多项式及一元有理函数的图像表示开始。由此得出的曲线与坐标轴的交点就是多项式的零点,自然而然地引导到方程的近似数值解。

(2) 曲线的几何图像自然而然地给微商和积分这两个概念提供了直观背景:曲线的斜率引进前者;曲线与 x 轴围成的面积导致后者。

(3) 积分过程(或就其本义来说是求面积过程),当其结果不能用有理函数及代数函数显式表示时,就会产生新的函数,所以这种新函数的引入是极其自然而前后统一的。例如,求双曲线面积过程,就定义了对数

$$\int_1^x \frac{dx}{x} = \log x;$$

而求圆面积过程可以很容易地归结为积分

$$\int_0^x \frac{dx}{\sqrt{1-x^2}} = \arcsin x;$$

即归结为反三角函数。你们知道,循此思路,前进一步就引入一类新的函数,特别是椭圆函数。

(4) 通过一个统一的原理即泰勒定理探讨函数的无穷幂级数的展开式。

(5) 将这个方法再向前推进,就得出柯西—黎曼的复变解析函数理论。它的基础是柯西—黎曼微分方程及柯西积分定理。如果我们想把根据以上的概括所得出的结果说得更确定一些,我们可以说,进程 A 是建立在把一门学科进行分解的概念上的,即把一个总体分成一系列互相独立的部分,使各个部分独立发展,尽量少借助于其他部分的知识,尽可能避免引入相邻领域的概念。进程 A 的理想是把各个局部领域的知识结晶为一个逻辑封闭系统。相反,进程 B 把主要重点放在各局部领域的有机结合上,放在各个局部的互相促进上,因而宁可采用统一的观点来理解好几个领域的方法。进程 B 的支持者的理想是把数学科学的总和理解为一个巨大的相互联系的整体。

人们不免要问这两个方法哪一个更有效?对于没有特殊的数学抽象天赋的学生,哪一种方法更好?要清楚地认识这一点,只要想一下函数 e^x 和 $\sin x$ 的例子。下面我们就从这些例子出发来谈函数 e^x 和 $\sin x$,在这方面以后还有许多可谈的内容。根据进程 A,不幸的是中学里几乎全是照此办理的,这两个函数的出现是完全不相干的。引入 e^x 或对数是为了数值计算的方便,但 $\sin x$ 却出现在三角形的几何讨论之中。这样怎么能使人理解两者之间所存在的如此简单的关系呢?尤其是因为两者一再出现的领域或与数值计算技术毫不相干,或与几何毫不相干,而是各有其趣,作为所讨论领域的规律的自然表现,这怎么能使人理解其相互关系呢?应用到 e^x 上的规律,有复利律或有机增长律,而一涉及振动问题就少不了用到 $\sin x$。从这些名称和情况就可以看出,两者的应用范围相去甚远!在进程 B 中,其中联系是以很容易理解的方式出现的,并与这些函数的意义相一致,这一点是从一开始就加以强调的。事实上,函数 e^x 和 $\sin x$ 在

这里是同出一源的，是从求简单曲线的面积产生的，由此很快就把人们引到最简单的微分方程

$$\frac{de^x}{dx}=e^x,\ \frac{d^2\sin x}{dx^2}=-\sin x。$$

这两个相应的微分方程当然是上述一切应用的基础，后面我们就可以看到这一点。

为了完整地理解数学的发展，我们还必须考虑到进程 C，它与进程 A，B 或者平行发展，或者含于其中，而且往往起到重要的作用。它涉及一种叫做算法的方法（algorithm——算法，是一个阿拉伯数学家的缩名）。一切有次序的形式运算归根结蒂都是一种算法。特别来说，字母运算是一种算法。我们一再强调，算法在科学发展过程中一直起着极为重要的作用。它作为半独立的力量，由公式本身规律推动其前进，并不取决于数学家的意图和认识，甚至往往相反。在无穷小分析初创时期，正如我们以后将看到的，算法往往强行推出新概念和新运算，后来才被人们所公认。甚至在数学发展水平到了很高的阶段时，算法研究还能起到很有效的作用，而且事实上也是如此，所以我们有理由称其为数学发展的基干。如果我们对这些情况不屑一顾，不过当作一些“形式上的”发展，像今天有时所作的那样，那就一定会完全抹杀历史。

现在让我通过数学的历史更仔细地探讨一下这些不同发展方向的对比，当然我只限于叙述最重要的发展特点。进程 A 和进程 B 之间的根本差别，将放到整个数学范围内来看，这样可以比前面看得更清楚一些，因为前面我们的思想只集中在分析上。

我们发现，在古希腊，纯数学和应用数学是截然分开的，这可以追溯到柏拉图（Plato）和亚里士多德（Aristoteles）。首先，无人不知的欧几里得几何结构属于纯数学的范围。在应用领域，古希腊人特别发展了一种所谓一般数的数值计算法。说实在的，一般数在当时并不受到重视，这种偏见至今仍很严重，当然主要在那些不会进行数

值计算的人。若说对一般数还有一点重视，可能由于它的发展结合三角和实际测量的需要，不过测量学之类在某些人看来还是不够高尚的。尽管如此，一般数的地位在人们眼里还是多少有了提高，可能是由于它在天文学中的应用，而天文学虽然同测量学有关，总是被认为是最高尚的学术之一。这几句话就可以使你们明白，希腊人的科学研究是有明确的分野的，各有严格的逻辑界限，完全是按进程 A 发展的。不过话又说回来，希腊人对于进程 B 范围内的思想见解也不是一无所知的，这些思想见解可能曾给他们以启发，或曾当作他们科学发现的最初依据，尽管在他们看来这些见解的最终表述，离不了进程 A 的形式。最近发现的阿基米德(Archimedes)手稿[①]，就非常明确地表明了这种情况。在那份手稿中，阿基米德通过力学研究进行了体积计算，他的计算方式是完全现代化的，看起来很顺眼，同严格的欧几里得系统截然不同。

除了希腊人外，特别值得一提的是印度人，他们在古代的时候，扮演了我们现代记数系统创造者的角色，而后来阿拉伯人又成了这种记数系统的传播者。首先用字母进行运算的也是印度人。这些进展显然属于进程 C 发展的范围，是一个算法发展过程。

谈到现代的情况，我们首先可以把数学的文艺复兴时代推到1500年左右，那一时期出现了整整一系列的伟大发现。作为一个例子，我要提到《三次方程的形式解》[卡丹诺(Cardano)公式]，收在卡丹诺所著1545年出版于纽伦堡的 *Ars Magna* 一书中。这是一本意义非常重大的著作，包含着现代代数的萌芽，远远超出了古代数学的框架。不过说真的，这个著作不是卡丹诺本人的著作，据说不仅他的著名的公式，而且其他内容，都取自他人。

1550年以后，三角计算进入了显著地位。第一部伟大的三角表的出现，是为了适应天文学的需要。谈到天文学，我只要提一下哥白

---

① 见《阿基米德文集》，剑桥大学出版社。

尼的名字就可以了。大约从 1600 年以后，对数的发明又成了这一发展的继续。苏格兰人纳皮尔(J. Neper)编了第一部对数表，实际上只包含三角函数的对数。由此我们可以看出，这 100 年间数学发展的道路是与进程 A 相对应的。

现在我们讲到了 17 世纪，并真正进入了现代，这时进程 B 清晰地占据了主导地位。1637 年出现了笛卡儿的《解析几何》，建立了数和空间之间的联系，为以后一切发展提供了基础。这本书后来有一本重印本，使我们读起来方便易懂。紧接着就接触到 17 世纪的两大问题——切线问题和求积问题，换言之，即微积分问题。就微积分的本来意义来说，当时所缺的知识就是还不知道这两个问题是密切相关的，是互为逆运算的。后来认识到了这个事实，这是 17 世纪末才取得的一个主要的伟大进展。

但是在此之前，在 17 世纪过程中，无穷级数，特别是幂级数的理论已经出现了，当然还不成其为今天代数分析的意义上的独立学科，而是同求积问题非常密切地结合着的。这个理论的开创人是尼柯拉斯·默卡多(Nicolaus Mercator，1620—1687 年，这是德国名字考夫曼的拉丁叫法，不是发明默卡多投影的那个默卡多)。他思想敏锐，通过除法，把分式 $\frac{1}{1+x}$ 变换成级数，然后对所得的级数逐项积分，以便求出 $\log(1+x)$ 的级数展开：

$$\log(1+x)=\int_0^x \frac{\mathrm{d}x}{1+x}=\int_0^x(1-x+x^2-x^3+\cdots)\mathrm{d}x$$
$$=x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\cdots。$$

这就是他的计算程序的实质，当然他还没有用我们这种简单的符号，如 $\int$，$\mathrm{d}x$，…，而是用了笨拙得多的表示式。在 17 世纪 60 年代，牛顿(1643—1727 年)接手做了下去，把这个计算用到了他在以前建立的一般二项式的级数上。在这个过程中，他用类比法得出了结论，即

把结论放在已知最简单的例子的基础上，而没有严格的证明，也不知道级数展开得以成立的范围。这里我们又看到了算法过程(进程 C)的作用。牛顿把二项式级数应用于 $\frac{1}{\sqrt{1-x^2}}=(1-x^2)^{-\frac{1}{2}}$，并利用默卡多的计算过程，得出了对于 $\int_0^x \frac{\mathrm{d}x}{\sqrt{1-x^2}}=\arcsin x$ 的级数。他对这个级数以及 $\log x$ 的级数进行了非常巧妙的反演，求得了 $\sin x$ 和 e^x 的级数。后来勃鲁克·泰勒(Brook Taylor，1685—1731 年)在 1715 年发表了《函数展开为幂级数的一般原理》，为这一系列数学发现作了总结。

前面已指出，17 世纪末无穷小演算的起源，归功于 G·W·莱布尼兹(1646—1716 年)和牛顿。牛顿的基本观念是流动的概念，变量 x，y 都被看做为时间 t 的函数 $\varphi(t)$，$\psi(t)$，随时间"流动"而流动。牛顿相应地称之为变流，并把我们所称的微分系数的 $\dot{x}$，$\dot{y}$ 命名为流数。你们看，这里的一切都是牢牢地建立在直觉的基础上的。

莱布尼兹的表达形式也是这样，他的第一部著作出版于 1684 年。他自己宣布说，他的最伟大的发现就是《一切自然现象中的连续性原理》，是这样的一句话"Nature non facit saltum"。他把数学的发展建立在这个概念上，这又是进程 B 的另一个例子。但是进程 C 算法的影响对于莱布尼兹也十分强烈。他给我们提供了在算法上很有价值的符号 $\frac{\mathrm{d}y}{\mathrm{d}x}$ 和 $\int f(x)\mathrm{d}x$。

从这个概念中可以看出，总的来说，17 世纪的伟大发现是属于进程 B 的发展范围。

到了 18 世纪，这个发现时期起先继续朝同一个方向发展。这里要提到的最杰出的人物是欧拉(1707—1783 年)和拉格朗日(1736—1813 年)。在这个时期，最广义的微分方程理论(包括变分法)得到了发展，解析几何和分析力学也发展了。到处都有可喜的进展，就像

在地理学方面一样，在发现美洲后，走遍了一些新的国家，并从各个方向进行了探测。但是当时还没有精密测量的概念，所以起先人们对于新地方的位置观念是完全错误的（哥伦布起先以为他到了东亚！）。在新征服的数学领域内也正是这样，人们所掌握的无穷小演算，起先离可靠的逻辑指导方向还很远。人们甚至对于无穷小演算与他所熟悉的老领域的关系抱有错觉，认为对无穷小演算无法进行逻辑分析，把它看成是某种神秘的东西。

当时理论所凭借的依据是多么的不可靠，这个情况到了打算编新的课本，清晰地介绍新的科目时就暴露了出来。那时明确了，按进程 B 去做再也不合适了，首先认识到这一点的就是欧拉。说实在的，他对于无穷小演算是没有严重怀疑的，但他觉得对初学者造成的困难和畏难情绪太多了。从这种教学法的角度来看，他以为最好有一个选修课程，如他在他所编的课本 *Intvoduction in analysin infinitovum*（1748 年，《无穷小分析引论》）所提供的那样。这个课程就是我们今天所说的代数分析。连带着，他特别把无穷级数和其他无穷过程理论包括了进去，这个理论后来又被他当作构造无穷小演算的一个基础。

差不多 50 年后，到了 1797 年，拉格朗日在 *Thèorie des fonctions analytiques*（《解析函数理论》）中采取的态度就激进得多了。他对当时的无穷小演算基础抱有严重怀疑，非要把当时的无穷小演算作一般的知识学科完全放弃，而仅视之为一堆应用于某些函数特例的形式法则，这样才能使他心安。确实，他仅只考虑那种可以用幂级数表示的函数

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots。$$

他把这些函数称作为解析函数，意指在分析中出现的、可加以合理应用的函数。然后对于函数 $f(x)$ 的微商，通过第二个幂级数纯形式地下了定义，正如我们以后会看到的那样。这样，微分和积分所研究的，就是幂级数的相互关系。仅限于形式考虑的这种观点，当然暂时

排除了许多困难。

正如你们所看到的，欧拉造成的转折，尤其是拉格朗日的全部方法，严格地属于进程 A 方向的发展，因为直观的发展被严格的、封闭的推理系统所代替了。他们这两个人对中学数学教学一直有深刻的影响，今天学校里不学习微积分本身，而学习无穷级数或根据待定系数方法通过幂级数来解方程，就表现出欧拉一书和拉格朗日思想对后代的影响。

现在我们来谈 19 世纪的情况，主要是通过“收敛性的判定”使高等数学分析有了一个比较可靠的基础。关于收敛性，在此以前也曾考虑过，不过考虑得很少。18 世纪是人们对收敛和发散蒙眬不分的“乐陶陶”的时期。甚至在欧拉的引论中，发散和收敛也相安无事地并列出现。但是在新的世纪开始的时候，高斯(1777—1855 年)和阿贝尔(N. Abel, 1802—1829 年)对收敛性作了第一批严格的研究；在 19 世纪 20 年代，柯西(1789—1857 年)又在他的讲演和著作中为现代意义的无穷小演算奠定了第一个严格的基础。他不仅像前人有时所做的那样，相应地用有限商和有限和作为微商及积分的精确定义，而且通过平均值定理第一次为无穷小演算建立了逻辑上相容的结构。这一点，我们以后还要回过来加以充分介绍。上述这些理论也具有进程 A 的性质，因为它们同其他知识部分隔开，并以逻辑的系统方式研究微积分。虽然这些理论有了彻底的改变，驱散了对于微积分的旧偏见，但对中学教学没有影响。

对于 19 世纪的进一步发展情况，现在只有一点地方还要加以强调。首先我要谈到进程 B 方向的若干进展，即根据柯西和黎曼理论，与复变函数论一起出现了现代几何和数学物理。首先从事这 3 大领域研究的带头人物都是法国人。这里也附带说一下数学表述的风格。在欧几里得几何中，一切问题都根据“假设，结论，证明”这几个步骤，有时还加上“讨论”，确定问题得以成立的界限。这种观念是很普遍的，以为数学叙述每次总是这 4 步。但是恰好在 19 世纪，兴

起了一种新的数学表述的艺术，特别是在法国数学家中。他们的这种手法可以称为艺术性推导手法。如蒙日(Monge)的著作，或最近出的一本书——皮卡(Picard)的《数学分析》，读起来都好像是一本写得很好的精彩的小说。这类风格适合进程 B 方法的需要，而欧几里得式的叙述方法实质上应列入进程 A 方法。

在这些方面取得突出成就的德国人中，我要提出雅可比(Jacobi, 1804—1851 年)、黎曼(1826—1866 年)以及稍后一点的克莱布什(1833—1872 年)，还有挪威人李(Lie, 1842—1899 年)。这些人本质上都是走进程 B 的方向，不过他们偶尔也表现出算法方式的痕迹。

19 世纪中期以后，进程 A 思想方法又因魏尔斯特拉斯(1815—1897 年)的出现而占据显著地位。他从 1856 年起开始在柏林从事教学。我已经拿魏尔斯特拉斯《函数论》作为进程 A 方法的一个例子。新近关于几何公理的研究，同样属于进程 A。完全是属于欧几里得方向的研究，在叙述方式上也很接近。

关于历史简介，就讲到这里为止。这里只能间接提到的许多观点，以后还可以更充分地加以讨论。总之，我们可以说，在过去几世纪的数学史上，上述两种主要研究方法都有重要地位，其中每一种方法——有时是两种方法交替地——都使数学这门科学得到了重大的进展。可以肯定，只要不忽视其中任一个方法，数学一定会始终如一地朝着这些方向发展，希望每一个数学家按最吸引自己的方向进行研究。

不过正如我已经指出的，中学的数学教学早已受到进程 A 方法的片面控制。因此，数学教学改革每走一步，都必须更多强调进程 B 方法。就这一方面来说，我首先想到引入直观启发的教学方法，更注重空间观念本身，特别是函数观念，使空间观念和数的观念融合起来！我的目的是通过这些讲座促成这种趋势，尤其是因为我们惯常参考的初等数学书，如韦伯—韦尔斯坦因、特罗夫克、M·西蒙的书，

几乎毫无例外地代表进程 A 的方法。在引论中，我已提请你们注意这种片面性。

好了，先生们，这些插话已经够了，下面就让我们讲本课程的下一个分支。

第二部分

代　　数

我先提几本代数教科书，以便向你们稍微介绍一下丰富的代数文献。首先要提出塞雷特(Serret)著《代数教程》[1]，这本书过去在德国用得很多，具有很多优点。但现在我们已经有了两本广泛采用的伟大的德国教科书：H·韦伯著《代数教科书》[2]及E·内托(E. Netto)著《代数课本》[3]。这两本书都是两卷本，极其完整地处理了代数学中最难的部分，并且对于广泛的专门研究也非常适用。我个人觉得，这两本书对于未来教师的一般需要来说过于详尽，也太贵了一点。比较适用的是G·鲍尔(G. Bauer)著《代数课本》[4]，这本书使用方便，极少超出教师应该掌握的范围。就实用方面来说，对方程的数值解，这本书还可以C·龙格(C. Runge)著的小册子《实用数学分析》[5]作补充，因此我也郑重推荐龙格的这本书。

现在转到比较窄的主题，请允许我说明一下，这里由于篇幅所限，我不能对代数作系统的介绍，只能侧重某一方面作有选择的讲解。如果我着重讲不幸被其他著作忽略的，然而有助于理解中学教学的部分，那真是再好也没有了。我的关于代数的全部讲法集中于一点，即方程图解法或一般地可称为几何直觉法的应用。单单这个领域就可以成为代数学中的一大章，它涉及许多问题。但即使这一部分也显然只能选出最重要和最有趣的东西来讲；不过在这样做的时候，我们将把各个相隔非常远的领域有机地组织起来，所以我们是按进程 B 的精神来研究数学的。第一步将处理实未知数方程，以便随后考虑复量。

---

① 塞雷特著，1866年巴黎第三版(1910年第六版)。

② 1898年/1899年布朗施魏格第二版(1924年R·弗里克(R Fricke)重新修订，第1卷)。

③ 内托，1896年/1899年莱比锡版。同时参阅克里斯特尔(Chrystal)《代数教科书》(两卷本)，麦克米伦出版公司版，以及M·博克(M. Böcher)著《高等代数引论》，麦克米伦公司出版。

④ 鲍尔著，1910年莱比锡第二版。并参阅E·内托著《初等代数》(1913年莱比锡第二版)，以及H·韦伯《代数教科书》，小卷本，1921年布朗施魏格第二次印刷。还可以参阅H·法因著《大学代数》，霍尔(Hall)及乃特(Knight)著《高等代数》，麦克米伦出版公司版。

⑤ 龙格著，1921年莱比锡第二版。并参阅H·桑登(H. V. Sanden)《实用数学分析》，达顿有限公司版。

第五章

含实未知数的实方程

§5.1　含一个参数的方程

我们从一个可以用几何方式处理的很简单的例子说起，这个例子就是未知数为 x 的实代数方程

$$f(x,\ \lambda)=0,$$

其中出现一个参数——λ。如果用第二个变量 y 代替 λ，把

$$f(x,\ y)=0$$

当作 x，y 平面上的一条曲线(图 5.1)，那么就能以最简单的方式得到一个几何表示：这条曲线与平行于 x 轴的直线 $y=\lambda$ 的交点，就是方程 $f(x,\ y)=0$ 的实根。当我们大致画出这条曲线(如果 f 不太复杂就很容易画)，再让平行线随 λ 的变化而改变位置，我们就能一眼看出实根数目的变化。当 f 关于 λ 是线性关系时，亦即对于方程

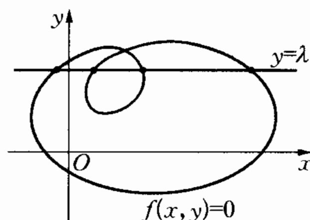


图 5.1

$$\varphi(x)-\lambda\psi(x)=0,$$

这种方法就特别有效。

如果 φ 和 ψ 是有理的，那么曲线 $y=\dfrac{\varphi(x)}{\psi(x)}$ 也同样会是有理的，

而且也很容易画出来。在这些情况下，往往可以很方便地利用这个方法来近似计算方程的根。

作为一个例子，请考虑下述二次方程

$$x^2 + ax - \lambda = 0。$$

曲线 $y = x^2 + ax$ 是一条抛物线。立即可以看出，当 λ 取什么值时，因抛物线与平行线相交于两点、相切于一点或不相交而使方程的实根个数相应地为 2，1 或 0(图 5.2)。我觉得，对中学高年级学生介绍这种简单而又一目了然的作图法是很适当的。

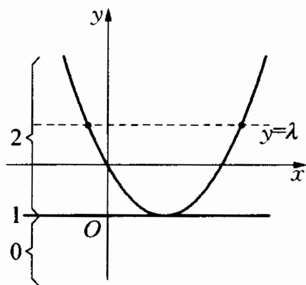


图 5.2

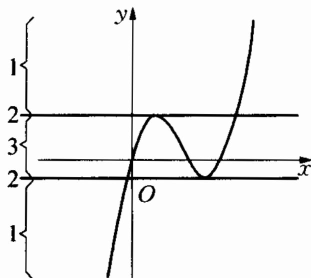


图 5.3

作为第二个例子，让我们取一个三次方程

$$x^3 + ax^2 + bx - \lambda = 0,$$

从中得出三次抛物线 $y = x^3 + ax^2 + bx$，其形状随 a，b 值的不同而不同。在图 5.3 中，假定 $x^2 + ax + b = 0$ 有两个实根。容易看出，平行线可分成两组，一组与曲线交于一点，另一组与曲线交于 3 点。这里可以有两个产生双重根的极限位置。

§5.2 含两个参数的方程

当一个方程内出现几个参数时，假如出现两个，用作图方式来处

理这类问题就需要更多的技巧,但所得的结果却更广泛、更有趣。下面以两个参数 λ,μ 线性地出现的情况为限,记 t 为方程中的未知数,问题为确定方程

$$\varphi(t)+\lambda\chi(t)+\mu\psi(t)=0 \tag{1}$$

的实数根,其中 φ, χ 和 ψ 是 t 的多项式。

如果 x, y 是通常的点的直角坐标,则 xy 平面上的每一条直线将由方程

$$y+ux+v=0 \tag{2}$$

给出。我们称 u, v 为直线坐标。这样,$-u$ 就是该直线与 x 轴夹角的正切,而 $-v$ 就是在 y 轴上的截距(图 5.4)。请把点和线看做同等重要,并对点坐标与线坐标给予同样的注意。这一点以后会认为是十分重要的。因此可以说,方程 $y+ux+v=0$ 指出了线(u, v)和点(x, y)的联结位置,即此点在该直线上,而直线通过该点。

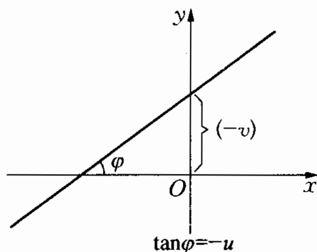


图 5.4

为了用几何方式来解释方程(1),请把方程(1)与方程(2)看成是等同的。可以采取两个根本不同的方式达到此目的。下面就分别加以考虑。

(1) 考虑方程

$$y=\frac{\varphi(t)}{\psi(t)},\quad x=\frac{\chi(t)}{\psi(t)}, \tag{3a}$$

$$u=\lambda,\qquad v=\mu。 \tag{3b}$$

如果 t 是参变量,则方程(3a)表示 xy 平面上的一条确定的有理曲线,称为方程(1)的正规曲线。因为曲线上的每一点对应于 t 的一个

确定值，据此即可以在曲线上确定某个 t 值的刻度。通过方程(3a)，可以按需要而计算出任意多的点，因而就可以把正规曲线的刻度按所希望的精确性画出来，例如在毫米纸上画出来。对 λ 和 μ 的每一对确定的数值，方程(3b)表示平面上的一条直线。由前述可知，方程(1)表明正规曲线上的点 t 位于这条直线上。因此，如果找出正规曲线与这条直线的一切实交点，读出这些实交点在曲线刻度上的参数值，就可以得出方程(1)的一切实根。正规曲线总由方程(1)的形式一次就确定下来，而不管参数 λ，μ 可能取什么样的特殊值。对于具有确定值的 λ，μ 的每一个方程，都有一条表示它的由上述方法确定的直线，使得平面上的一切直线一般地说都起到前节所仅用到的水平直线那样的作用。

作为一个例子，取二次方程

$$t^2+\lambda t+\mu=0,$$

这里的正规曲线由下列方程

$$y=t^2,\quad x=t \quad 或 \quad y=x^2$$

给出，即正规曲线是图 5.5 所示的具有所标刻度的抛物线。可以立刻从它与直线 $y+\lambda x+\mu=0$ 的交点读出方程的实根。图 5.5 还表明，方程 $t^2-t-1=0$ 的两个根相应地处在 $\frac{3}{2}$ 和 2 及 $-\frac{1}{2}$ 和 -1 之间。这个方法与前节的方法相比，其主要优点是，如果利用平面上的一切直线的话，则用同样的抛物线就能解出所有二次方程。因此，如果想求许多二次方程的近似解，就可以非常有效地利用这个方法。

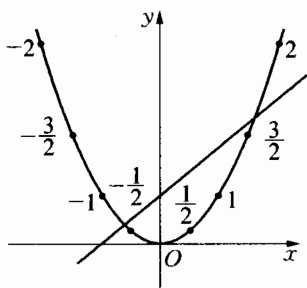


图 5.5

一切三次方程也可用类似的方法处理。通过线性变换，可将其化为简化形式：

$$t^3+\lambda t+\mu=0。$$

这里的正规曲线是立方抛物线

$$y=t^3,\ x=t\quad 或\ y=x^3,$$

如图 5.6 所示。这个方法我觉得也可以在中学里用，学生一定会从画那样的曲线中得到乐趣。

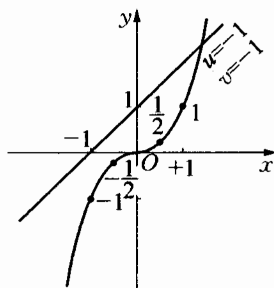


图 5.6

(2) 解释方程(1)的第二个方法是从第一个方法中得出的，它应用对偶原理，即让点坐标和线坐标互换位置。为此，让我们把方程(2)中各项的次序颠倒一下，记为

$$v+xu+y=0,$$

并置

$$v=\frac{\varphi(t)}{\psi(t)},\quad u=\frac{\chi(t)}{\psi(t)},\tag{4a}$$

$$x=\lambda,\qquad y=\mu,\tag{4b}$$

使之从形式上与方程(1)等同。

如果取 t 为参数，则方程(4a)就表示包络一条确定曲线的直线族，据新的解释，即包络了方程(1)的正规曲线。它是有理类曲线，因为它在线坐标上用一个参数的有理函数表示。每一条切线以及相应的切点，都由一个确定的 t 值来决定。从而再一次得到正规曲线上的刻度。据方程(4a)，画出足够多的切线，就可以把曲线和刻度都按所要求的精确度画出来。借助于方程(4b)，每一对参数 λ, μ，可确定 xy 平面上一点。据方程(1)，正规曲线(4a)上对应 t 的切线必定通过此点。因此，读出正规曲线上通过点 $x=\lambda$, $y=\mu$ 的所有切线所对应的参数值 t，即可求得方程(1)的所有实根。和前面一样，正规

曲线全由方程(1)的形式所确定。对于参数 λ，μ 的给定值，这种形式的每一个方程，都由平面上的某一点表示，或由其相对于曲线的位置来确定。

现用与前面一样的例子来说明。对应于二次方程

$$t^2+\lambda t+\mu=0,$$

正规曲线将是以下直线族的包络：

$$v=t^2,\quad u=t。$$

此包络是一个顶点在原点的抛物线。作在细横格纸上的图立刻表明，$t^2+\lambda t+\mu=0$ 的实根就是由点 λ，μ 至抛物线的切线的参数 t(图 5.7)。

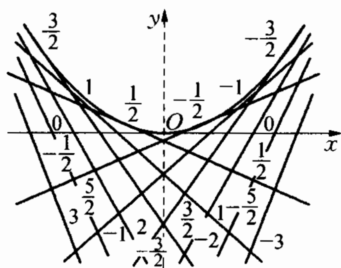


图 5.7

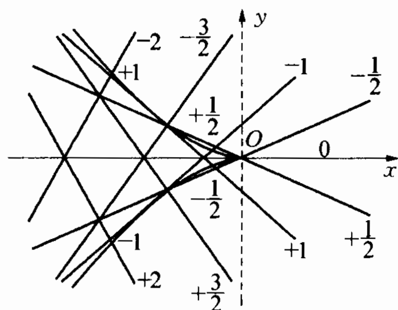


图 5.8

对于三次方程

$$t^3+\lambda t+\mu=0,$$

正规曲线

$$v=t^3,\quad u=t$$

为在原点有一个尖点的第三类曲线，如图 5.8 所示。

这个方法也可以换一个略微不同的方式来说明。如果考虑所谓三项方程

$$t^m + \lambda t^n + \mu = 0,$$

那么就可以借助于下列参数方程

$$f(t) = t^m + xt^n + y = 0$$

来表示正规曲线的切线族。要想求出正规曲线的直角坐标方程，可以像通常那样，从最后一个方程和对 t 求导数而得方程

$$f'(t) = mt^{m-1} + nxt^{n-1} = 0$$

之间消去参数 t。因为作为一族直线的包络的正规曲线，是各条直线与相邻直线（分别对应于 t 和 $t + \mathrm{d}t$）的交点的轨迹。如果不消去 t，把 x 和 y 表示为 t 的函数，则得

$$x = -\frac{m}{n}t^{m-n}, \quad y = \frac{m-n}{n}t^m, \tag{5a}$$

这就是正规曲线的点方程。

上面选用了二次方程和三次方程作为例子，按此途径相应地求得这些方程的正规曲线为

$$x = -2t, \quad y = t^2,$$
$$x = -3t^2, \quad y = 2t^3。$$

这些曲线就是图 5.7 和图 5.8 所画的曲线。

要强调一下，实际应用这个方法的是 C · 龙格，他在讲稿和练习中都曾用这个方法，并证明了这个方法对实际解方程特别适用。这些作图方法，可以在中学教学中任选一种来用，用起来都有好处。

如果把已提出的两种方法互相比较一下，就会发现第二种方法有明显的优点，至少符合一个确定的、非常重要的目的，即可以把一切有给定个数的实根的、确定类型的方程形象地表示出来，因为根据第一种方法一切方程是用一组直线来表示的；而根据第二种方法，却是用点的区域来表示的。由于几何直觉，或者说根据习惯，第二种方法本质上要比第一种方法更容易掌握一些。

下面我马上就要通过二次方程的例子来说明按这种解释可得出怎样的结果(图 5.9)。从抛物线外侧的各点到曲线,可作两条切线;而从抛物线内侧各点,不能作任何切线。因此,这两个区域相应地表示带有两个根的方程和不带根的方程族。对在抛物线上的各点,仅存在一条单一的切线,但可以计算两次,因而正规曲线本身在一般情况下就是那些点的轨迹,其坐标 λ, μ 可产生两个等根的方程,可称此轨迹为判别曲线。

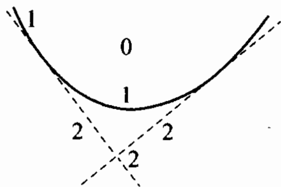


图 5.9

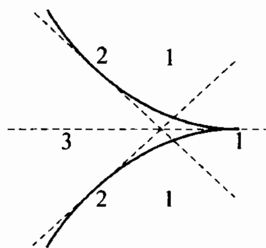


图 5.10

在三次方程的情况下,我们可以看到,从正规曲线尖角内的点到曲线,可以作 3 条切线。对于中线上的点,这是显而易见的,因为具有对称关系。而当点连续变化时,如果不穿过曲线,其切线数不可能改变。如果点(x, y)移到曲线上,则两条切线相重合。如果点移到曲线尖角的外侧,则两切线皆变为虚切线,只剩下一条实切线。因此,正规曲线尖角内的区域所表示的,是含有 3 个不同实根的全部三次方程;角外的区域所表示的,是只含有一个实根的方程;而对应于曲线上各点的,是含有一个单实根及一个二重实根的方程。最后,有一条三重切线穿过尖点,对应于含有一个三重根的方程 $t^3=0$。只要看一下图 5.10 就可以明白了。

如果按代数中的习惯,对根加上确定的限制,特别,如果想求给定区间 $t_1 \leqslant t \leqslant t_2$ 内的一切实根时,那么图 5.9 和图 5.10 就更加有趣,能说明更多的问题。如你们所知,斯图姆(Sturm)定理给出了这

个问题的一般解答。不过我们可以轻而易举地把图画完，就使这个一般问题得到一个令人满意的图像解答。为此目的，只消在正规曲线上作出由参数值 t_1，t_2 所决定的切线，并考虑平面上由这些切线所造成的分割情况。

为了把这些想法运用到二次方程上去，必须确定与 t_1 和 t_2 之间的抛物线弧相切的切线数。从抛物线这段弧与上述两条切线围成的三角形内的每一点，显然有两条切线通过（图 5.11）。如果此点横穿任一切线 t_1，t_2，那么两切线之一就会切于（t_1，t_2）弧以外的抛物线，它对我们就没有意义了。以抛物线和切线 t_1，t_2 为界的两个月牙形区域以内的点的切线，将只与弧（t_1，t_2）以外的抛物线相切；过抛物线内侧的点，则完全没有实切线。因此，对应于 $t \leqslant t_1$ 和 $t \geqslant t_2$ 的两个抛物线弧，对于我们想作的分割平面是没有意义的。这样就只剩下图中以实线标出的那样线，这些线以及画在它们上面的数字，使我们一看就可以对二次方程的全体在 t_1 和 t_2 之间有两个、一个或 0 个实根得到准确的答案。

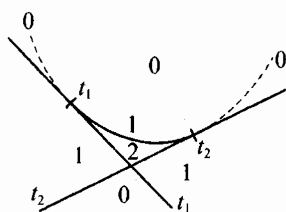


图 5.11

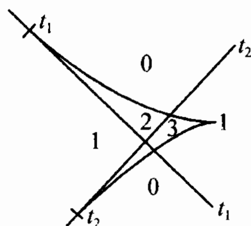


图 5.12

可以按类似的方法处理三次方程（图 5.12）。举例来说，取 $t_1 > 0$，$t_2 < 0$。作相应于这些参数值的切线，并且观察由这些切线及处于 t_1 和 t_2 之间的正规曲线的弧所造成的平面分割。通过尖点处四角形区域中的每一点，有 3 条实切线在 t_1 和 t_2 之间与曲线相切。如果点穿过切线 t_1，t_2 之中的任一条，就失去一条具有这种性质的切线。如果点穿过正规曲线，则两条切线均告失去。根据这些

考虑，可得出在 t_1 和 t_2 之间具有 3 个、两个、一个根或不带根的三次方程的平面区域图，如图 5.12 所示。为了了解这种几何作图法的巨大用处，只消试一试，如果用抽象的方式来考虑这种三次方程的分类，而不求助于任何空间直觉，那么所需时间之多就不成比例了。这种只要看一下图形就很明了的事，证明起来却不容易。

至于这种几何方法与众所周知的斯图姆、卡德修斯(Cartesius)和布登(Budan)-傅里叶(Fourier)代数准则之间的关系我只指出一点，就我们所考虑的种种方程而言，这种几何方法包括了它们的一切内容。你们可以在我的文章"Geometrische zur Abzählungder Wurzeln algebraischer Gleichungen"[①]及 W·戴克斯(W. Dycks)的"Katalog mathematischer Modelle"[②]中找到对这些关系的更充分的说明。我很高兴利用这个机会向你们介绍戴克斯的那个一览表，它是在德国数学会举办的 1893 年慕尼黑展览之际公布的，至今仍为数学模型方面最好的指导。

§5.3 含3个参数 λ, μ, ν 的方程

最后，我还要向你们说明一下，可以用类似的方法去处理含 3 个参数的方程。我们需用三维空间来代替平面。只需考虑特殊的四项方程就足够了。

$$t^p + \lambda t^m + \mu t^n + \nu = 0。\tag{1}$$

可以直接把这个方法运用于其他形式的方程。

除此方程，还要用到空间几何的一个条件，即一个点(x, y, z)和一个坐标为(u, v, w)的平面处于"统一的位置"，即这个平面$(u,$

① 重印于克莱因 *Gessamelte Mothematische Abhandlungen* 第 2 卷中，第 198—208 页。

② 参阅《教学及数学物理模型教具一览表》(慕尼黑，1892 年)，以及此表之补充部分(慕尼黑，1893 年)。

v, w)包含点(x, y, z)的条件是

$$z + ux + vy + w = 0, \tag{2}$$

或

$$w + xu + yv + z = 0。 \tag{3}$$

令记成方程(2)或方程(3)形式的这个方程恒等于方程(1),就可以完全像前面一样得出两个互相对偶的解释。

接着令

$$z = t^p, \quad x = t^m, \quad y = t^n。 \tag{2a}$$

这些方程确定了某一空间曲线,即四项方程(1)的正规曲线,同时也确定了曲线上值 t 的刻度。然后,我们考虑方程(1)中由系数 λ, μ, ν 所确定的平面

$$u = \lambda, \quad v = \mu, \quad w = \nu。 \tag{2b}$$

这样,根据方程(1),所述方程的实根就等于正规曲线(2a)与平面(2b)的实交点的参数值 t。

如果选择与前述方法相对偶的方法,就必须置

$$w = t^p, \quad u = t^m, \quad v = t^n。 \tag{3a}$$

这些方程所表示的是以 t 为参数的单参数平面族,可看做是空间曲线的密切平面族。此曲线和前面讲的一样,上面有 t 的刻度。这里所讲的曲线,是一种平面坐标表示的正规"类曲线",以区别前面用点坐标给出的正规"阶曲线"。如果现在结合第一种曲线,也考虑到点

$$x = \lambda, \quad y = \mu, \quad z = \nu, \tag{3b}$$

即知方程(1)的实根就是正规类曲线(3a)上的经过点(3b)的密切平面的参数值 t。

下面通过具体例子来说明这两种解释。我们收藏有一些模型,可以用来说明这两种解释,现在我拿给你们看。

斯图加特的 R·密姆克(R. Mehmke)用第一种方法制造了一种求方程数值解的仪器。他的模型是一个铜架(图 5.13),架子上有 3 根带刻度的垂直杆,可以装入一个弯曲的模板,用以表示三次、四次或五次方程(在简化为四项后)的正规曲线。不过要注意,我们的说明是以通常的直角坐标系为假设前提的,而密姆克则是这样确定坐标系的:他选了平面方程(2)的适当平面坐标,使系数 u, v, w,正好是这个平面在 3 根垂直杆刻度上的截距,此刻度可在杆上读出。为了确定一个平面 $u=\lambda$, $v=\mu$, $w=\nu$,在杆 w 上刻度读数为 ν 处搞了一个透视孔,再用一根线,把杆 u 上和杆 v 上的相应刻度读数连接起来。连接透视孔和拉线的射线族,就是我们所说的平面。通过透视孔,就能直观地看到此平面与正规曲线的交点,即此模板与此线的视交点。它们的参数值,即欲求的方程的根,可以同时从附着在模板上的正规曲线的刻度读出。这个教具的实用价值当然要看机械结构做得精巧不精巧,但人的眼力究竟有限,所以这种教具做得再好,其实用价值也是令人怀疑的。

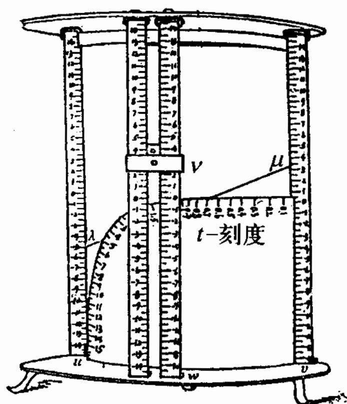


图 5.13

哈登斯坦因(Hartenstein)为了国家考试提出来的模型用的是第二种方法。它适用于所谓化约的四次方程,即

$$t^4+\lambda t^2+\mu t+\nu=0。$$

但每一个四次方程都可以简化为这种形式。我对于这个方法的讲法,将与我对于带两个参数的方程的讲法稍有不同。对于现在这个例子,我们必须考虑单参数的平面族,其平面坐标由方程(3a)给出,而其点方程则记为下面形式

$$f(t) = t^4 + xt^2 + yt + z = 0。 \tag{4}$$

这些平面的包络是平面 $f(t)=0$ 与其相邻平面 $f(t+\mathrm{d}t)=0$ 相交的直线族(即通过消去 $f(t)=0$ 及 $f'(t)=0$ 之间的 t 所得方程)的可展曲面。但是为了求得正规曲线,必须找出平面族的密切图形,即 3 个相继平面相交的点轨迹。正如你们所知道的,这个轨迹是那个可展曲面的尖棱,它的坐标作为 t 的函数,是从 3 个方程 $f(t)=0$, $f'(t)=0$, $f''(t)=0$ 中求出的。这个例子中的 3 个方程是

$$t^4 + xt^2 + yt + z = 0,$$
$$4t^3 + x \cdot 2t + y = 0,$$
$$12t^2 + x \cdot 2 = 0。$$

由此可求出

$$x = -6t^2, \quad y = 8t^3, \quad z = -3t^4。 \tag{5}$$

这些式子所表示的,是方程(4)的正规类曲线的点方程。据方程(3a),此曲线的平面方程可记为

$$w = t^4, \quad u = t^2, \quad v = t。 \tag{6}$$

两个形式对 t 都是四次的,因而正规曲线既是四阶曲线,也是四类曲线。

为了更详细地进行研究,我们考虑包含该曲线的几个简单曲面。首先,(5)式对任何 t 均满足方程

$$z + \frac{x^2}{12} = 0,$$

因而正规曲线处于一个二次抛物柱面上,柱面的母线平行于 y 轴。同样,我们有关系

$$\frac{y^2}{8} + \frac{x^3}{27} = 0,$$

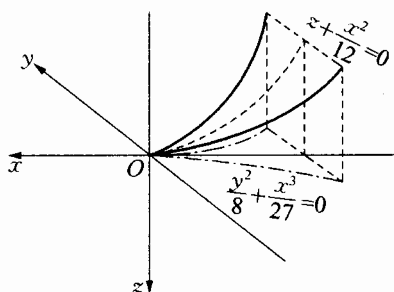


图 5.14

因而母线平行 z 轴的这个三次柱面也包含正规曲线。此外，正规曲线是这两个柱面的有限交。记住这些，就可以对正规曲线的走向有一个近似概念。这是一条挠曲线，对称于 xz 平面，在原点处有一个尖点（图 5.14）。

二次曲面

$$\frac{x\cdot z}{6}-\frac{3y^2}{64}=0$$

也包含该正规曲线，因为据(5)式，这个方程对 t 也恒满足。由此方程及三次柱面方程得出另一个线性组合，表示一个极其重要的包含正规曲线的三次曲面

$$\frac{xz}{6}-\frac{y^2}{16}-\frac{x^3}{216}=0。$$

现在考虑尖棱是正规曲线的可展曲面，这种曲面可以定义为正规曲线的切线全体。任何空间曲线

$$x=\varphi(t),\quad y=\psi(t),\quad z=\chi(t)$$

在点 t 处的切线由下列方程

$$x=\varphi(t)+\rho\varphi'(t),\ y=\psi(t)+\rho\psi'(t),\ z=\chi(t)+\rho\chi'(t)$$

给出，其中 ρ 是参数，这是因为曲线的切线的方向余弦与坐标对 t 的导数成比例。如果把 t 也当作变量，则上述方程中有两个参数 t, ρ 而成为可展曲面的参数表示式。这一切都是从众所周知的空间几何定理推导出来的。对于曲线(5)，可得出下面的可展曲面方程。如果用坐标(X, Y, Z)来表示该曲面，以区别于曲线的坐标(x, y, z)，则可展曲面方程为

$$X = -6(t^2 + 2\rho t),$$
$$Y = 8(t^3 + 3\rho t^2), \qquad (7)$$
$$Z = -3(t^4 + 4\rho t^3)。$$

这个曲面就是被拉紧的线所表示的直线族,即哈登斯坦因模型的基础(图 5.15)。

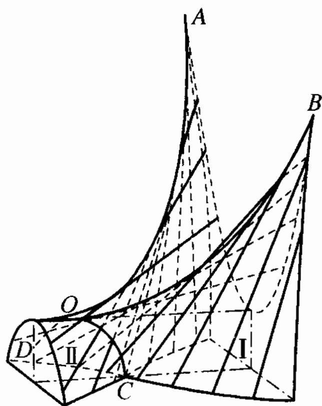


图 5.15

这种参数表示法为讨论曲面及进行曲面作图提供了最好的出发点。事实上,总想要研究曲面自身的方程,仅仅是出于习惯的压力。这个方程只要去掉(7)中的参数 ρ 和 t 即可得出。下面提供一个最简单的求方程步骤,其中有几个步骤的内在意义就不详谈。由(7)式可得出组合

$$Z + \frac{X^2}{12} = 12\rho^2 t^2,$$

$$\frac{X \cdot Z}{6} - \frac{Y^2}{16} - \frac{X^3}{216} = 8\rho^3 t^3。$$

它们每个都在曲线本身上等于 0(因为 $\rho = 0$)。如果使这些组合等于零,则得到上述通过曲线的两个曲面。从这些方程中去掉乘积 ρt,则得可展曲面的方程为

$$\left(Z + \frac{X^2}{12}\right)^3 - 27\left(\frac{X \cdot Z}{6} - \frac{Y^2}{16} - \frac{X^3}{216}\right) = 0。$$

因此,这个曲面是六次的,但它是由无穷远平面和一个五次曲面组合而成。

至于这个公式的意义,我要作以下的说明,供了解这个问题的人参考。这两个括号里的式子是四次方程

$$t^4 + Xt^2 + Yt + Z = 0$$

的不变量，上面就是以此作为出发点的。这些不变量在椭圆函数理论中起着重要的作用，在那里，通常记为 g_2 和 g_3。如你们所知，曲面方程的左边 $\Delta = g_2^3 - 27g_3^2$ 是四次方程的判别式，它变为 0 就表示出现重根。因此，我们这个可展曲面是四次方程的判别曲面，即有重根的点的集合。

作了这些理论解释之后，为此曲面作拉线模型就没有很大困难了。例如，通过参数方程(7)，可以确定这样一些点，在这些点处，我们所要表示的那些切线贯穿某些固定平面。这些平面可用木板或硬纸板做成，然后再在这些平面之间拉线。但要把模型做得真正漂亮、实用，把曲面和尖棱的全部走向表现出来，像前面所看到的那样，就需要反复试做，要有高超的技巧。图 5.15 就是把线拉直后所出现的曲面概貌，其中 AOB 是尖棱(参看图 5.14)。

在模型上可以看到一条双重曲线(COD)，沿着这条曲线有两叶曲面相交，这条曲线就是 XZ 平面上的抛物线

$$Y = 0, \quad Z - \frac{X^2}{4} = 0。$$

但这条抛物线上只有一半(即对 $X < 0$ 的一半)是实叶的交线，而另一半却孤立地位于空间中。对于习惯于用具体几何表示法来说明代数曲面理论的人，这种现象是不足为怪的。双重曲线的实枝既可作为实叶的交线，又可有一部分呈孤立状态，这是普遍现象。在后一种情况下，我们视之为曲面上虚叶的实交线。平面上的相应现象，了解的人比较多。大家都知道，在那种情况下，除了代数曲线的一般双重点(作为曲线上实枝交点而出现)以外，还有明显的孤立双重点，可以被看做虚枝的交点。

现在我们要详细搞清楚这种以正规曲线为尖棱的曲面有什么用。我们把正规曲线看做是有刻度的，或说得更确切些，对每一条切线都附以参数值 t，即该切线的切点的参数值。如果现在有人提出一个带确定系数(x, y, z)的四次方程，就只要求正规曲线的穿过点

(x, y, z)的密切平面，或使判别曲面的切平面通过该点，两者都一样。这样，实根或者就是与曲线接触点的参数值，或者就是相应切平面的参数值，视情况而定。出于密切平面在同曲线接触的地方与曲线相切，所以密切平面与曲线的每一个接触点，由点(x, y, z)看起来就是一个拐点，反之亦然。因而四次方程的实根，归根结底，就是从空间内的点(x, y, z)看去所见到的正规曲线的拐点的参数值t。

没有受过训练的眼睛，当然难以从模型上看清密切平面或曲线的这种拐点。但模型可以使人一目了然地看清另一件重要的事情，即根据实根数目对一切四次方程的分类。请通过对方程的抽象观察，看看到底会遇到哪些情况。如果α, β, γ, δ是实四次方程(4)的4个根，因为t^3的系数为0，故$\alpha+\beta+\gamma+\delta=0$。就实根来说，可能存在下列主要情况：

Ⅰ. 4个实根。

Ⅱ. 两实根和两共轭复根。

Ⅲ. 无实根，但有两对共轭复根。

如果现在提出属于类Ⅰ的两个方程，其根分别为α, β, γ, δ和α', β', γ', δ'，那么当然可以在保持其和为0的情况下，相应地把α, β, γ, δ连续转换为α', β', γ', δ'。与此同时，一个方程可始终通过同一类方程(即类型Ⅰ的所有方程)而连续地变化为另一个方程，从而构成一个连通的连续统。另两类方程也是如此。因此，我们的模型必须表现出空间被隔成3个连通部分，使每一部分中的所有点对应于各类方程的情况。

现在考虑这3类情况之间的过渡情况(见图5.16)。类Ⅰ是通过有两个不同实根和一个重(两个叠合的)实根的方程(记为2+(2))而转换成类Ⅱ的。同样地，在类Ⅱ和类Ⅲ之间有一个实双重根和两个复根的过渡情况，记为(2)。在我们的模型中，对应于这两类情况的必为判别曲面区域，此区域事实上把所有具有重根的方程都表现了出来。基于类似的考虑可以知道，对应于上述每一类方程，必

有一个属于此曲面的连通区域。于是 2+(2)和(2)这两类情况又可以通过带两个实双重根(记为(2)+(2))的情况而相互转换。两对根由此变成叠合的点,必同时属于判别曲面的两叶,即属于双重曲线的非孤立分支。相应地,判别曲面由一个双重曲线的一枝分成两部分,其中之一 2+(2)把空间区域Ⅰ和Ⅱ隔开;另一部分(2)把空间区域Ⅱ和Ⅲ隔开。再来看看正规曲线所处的情况。可以指出,由于它作为尖棱的性质,3 个切平面必在它的每一点上合而为一(密切平面),这就产生了一个三重根和一个单实根的情况:1+(3),这种情况只有在单根之一等于二重根时才会发生。因此,尖棱必完全处在曲面的第一部分,即2+(2)处。在棱的尖点 ($x=y=z=0$),有一个四重实根,可通过两个双重根,即(2)+(2)的情况重合而成。实际上,尖棱的尖点 O 也处在双重曲线上。至于双重曲线的孤立分支,则完全处于空间区域Ⅲ中,其上两对共轭复根合而为两个复重根。这两个双重根当然是彼此共轭的。

以上所列举的各种情况,在模型上都可以看到。在图 5.15 上,双重曲线右边的曲面内部是区域Ⅰ,左边是区域Ⅲ,外部是区域Ⅱ。在图 5.16[①] 中列出了实根的个数和相重数(对应于各部分空间、曲面、曲线区域的点),很容易完全搞清楚。在这幅图上,不在括号内的数字表示单实根数,其他数字像前面一样表示重根的相重数。

	Ⅰ		Ⅱ		Ⅲ
区　　域:	4		2		0
判别曲面:		2+(2)		(2)	
正规曲线:		1+(3)			
双重曲线:			(2)+(2)		(2 虚,重根)
尖　　点:			(4)		

图 5.16

① 图号为编辑所加。

第六章

复数域方程

现在来探讨复数域方程，而不再限于实数域。当然，我们还是只着重，而且比别处更着重可以用几何方法表示的东西。我们现在就从代数的最重要的定理说起！

§6.1 代数的基本定理

正如你们所知道的，这个定理就是：复数域的每一个 n 次代数方程一般有 n 个根，或者更准确地说，每一个 n 次多项式 $f(z)$ 可以分解成 n 个线性因子。

这个定理的所有证明都从根本上应用了复数 $x+\mathrm{i}y$ 的几何解释。我可以向你们提供高斯的第一个证明(1799 年)的思路，这个思路是完全可以用几何法来表示的。不过，高斯原来的讲法同这里的讲法稍有出入。

给出多项式

$$f(z) = z^n + a_1 z^{n-1} + \cdots + a_n,$$

可把它记作

$$f(x+\mathrm{i}y) = u(x,\ y) + \mathrm{i} \cdot v(x,\ y),$$

其中 u, v 是两个实变量 x, y 的实多项式。高斯的证明主要是考虑到了 xy 平面上的两条曲线

$$u(x,\ y)=0,\quad v(x,\ y)=0,$$

并证明这两条曲线至少有一个公共点。对此点，于是有 $f(x+\mathrm{i}y)=0$，即证明方程 $f=0$ 有第一个根。为此，只需研究此两曲线在无限大时，即离原点任意大距离时的情况。

如果 z 的绝对值 r 非常大，与 z^n 相比较，就可把 $f(z)$ 内含 z 较低幂的项忽略不计。如果在 x, y 平面上引入极坐标 r, φ，即令

$$z=r(\cos\varphi+\mathrm{i}\sin\varphi),$$

则据棣莫弗(De Moivre)公式，有

$$z^n=r^n(\cos n\varphi+\mathrm{i}\sin n\varphi)。$$

随着 z 的绝对值增大，被 $f(z)$ 逐渐逼近这个式子。由此立刻得出，u 和 v 相应地逐渐逼近函数

$$r^n\cos n\varphi,\quad r^n\sin n\varphi$$

因此，曲线 $u=0$, $v=0$ 在无穷远处的最终走向可分别由下列方程

$$\cos n\varphi=0,\quad \sin n\varphi=0$$

近似地给出。曲线 $\sin n\varphi=0$ 由 n 条穿过原点并与 x 轴成 0, $\frac{\pi}{n}$, $\frac{2\pi}{n}$, $\cdots$, $\frac{(n-1)\pi}{n}$ 度角的直线构成，而 $\cos n\varphi=0$ 则由 n 条穿过原点把这些角等分的射线构成(图 6.1 所示是 $n=3$ 的情形)。在图的中心部分，真的曲线 $u=0$, $v=0$，当然可能根本不同于这些直线，但必须随着远离原点而逐渐逼近这些直线。我们可以这样画出曲线走向的示意图来，即在一个大圆以外，保持这些直线不变，而在圆内则用任意曲线来代替它们(图 6.2)。但不管圆内曲线的情况如何，可以肯定，如果让围绕原点的圆足够大，则圆外的 u, v 的各分支必互相交替。从图形上可以看出，这些分支必在

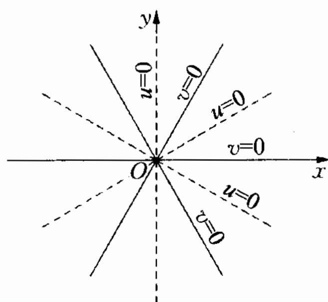


图 6.1

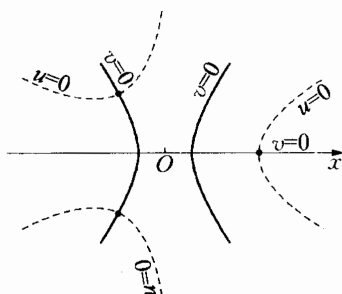


图 6.2

圆内相交。事实上，如果利用曲线的连续性质，就可以对这个判断给出严格的证明①，而这也就是高斯证明的实质。前面的论据是高斯证明思路的要点。如果找出一个这样的根，就可以分离出一个线性因子，于是我们又可以重复这个推理过程于另一个 $n-1$ 次多项式因子，依此继续下去，最后可以把 $f(x)$ 分解为 n 个线性因子，即可以证明 n 个零点的存在。

如果用一个特例把这个作法贯彻到底，这种推理方法就会清楚得多。一个简单的例子是

$$f(z) = z^3 - 1 = 0。$$

这种情况下显然有

$$u = r^3\cos 3\varphi - 1, \quad v = r^3\sin 3\varphi。$$

故 $v=0$ 仅由 3 条直线构成，而 $u=0$ 有 3 条双曲线状的分支。图 6.3 画出

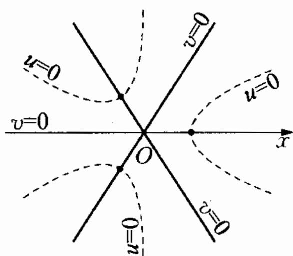


图 6.3

① 这里应该说明，高斯也不是完全没有考虑到几何方面。他的论文中所考虑的证明，首先由 A·奥斯特洛夫斯基(Ostrowski)使之算术化(*Göttingen Nachrichten*，1920 年，或《高斯科学传记材料》第 8 卷，1920 年)。有一段历史上有趣的事：这个基本定理的第一个证明是达朗贝尔(D'Alembert)得出的。实际上，他的证明中有一个错误，高斯注意到这一点，即达朗贝尔没有把函数的上极限与函数的极大值区别开来，并假设复变函数存在极限时就能达到它的上极限。这个论断一般不成立。

了这两条曲线的 3 个交点，也就给出这个方程的 3 个根。我认真地建议你们一定要拿几个较复杂的例子并计算到底。

关于上述基本定理，简单讲到这里就够了，因为我们不是在上代数课。让我在结束这一节时指出，使复数进入代数的意义在于使这个基本定理得到一个统一的说法。如果仅限于实数，那么只能说 n 次方程有 n 个根或少一些根，或完全没有根。

§6.2 含一个复参数的方程

现在就把用来讲代数的余下的时间用于以几何作图方法讨论复数方程的一切根(包括复根)，就像前面讨论实数方程的实根一样，不过我们只讨论含一个复参数的方程，并进一步假设这个参数仅线性地出现。研究一种简单的保角变换，就可以达到这一切要求。

设 $z=x+\mathrm{i}y$ 是一个未知数，$w=u+\mathrm{i}v$ 是参数，则所要考虑的那一类方程具有形式

$$\varphi(z)-w\cdot\psi(z)=0, \tag{1}$$

其中 φ, ψ 是 z 的多项式。设 n 为所出现 z 的最高次幂。根据基本定理，这个方程对 w 的每一个确定值恰有 n 个根 z，它们一般不相同。不过，相反地，由(1)式可知

$$w=\frac{\varphi(z)}{\psi(z)}, \tag{2}$$

即 w 是 z 的单值有理函数，并称之为 n 次的。如果干脆利用此函数在 z 平面与 w 平面之间所建立的保角变换，以此作为方程(1)的几何等价式，那么 z 作为 w 函数的多值性，从直观上看就不清楚。有一个补救的办法，就是用函数论中的常用办法，把 w 平面看做是由 n 个互叠的叶构成的，这 n 个叶再通过适当方法，借助于分支的切口而连接合成为一个 n 叶黎曼曲面。这些代数函数论中的曲面，都是你们所熟悉

的。这样,我们的函数就在 w 平面上的 n 叶黎曼面的点和单叶的 z 平面的点之间建立了一一对应关系,这种关系一般是保角的。

在详细研究这个变换之前,先确定某些约定是有帮助的,这些约定可以排除 w 和 z 的无穷远点所起的例外作用,而这种作用与这里的问题本质是无关的,这样我们能够用一般形式来叙述定理。由于这些约定还没有得到所应得到的广泛应用,请允许我多解释几句。这里不能把复平面上有一个无穷远点当作一种符号,因为从这种抽象概念中得不到适当的具体形象,所以必须求助于种种特殊的考虑和约定,使得一个有限点所具有的某个确定性质,对一个无穷远点也能说出究竟是什么。不过,如果彻底地用黎曼球面来代替复数图示的高斯平面,就可以达到一切要求。为此,只需设想直径为 1 且与 xy 平面相切的球面,把它的南极 S 置于原点,从北极 N 把平面作球极投影于球面上(图 6.4)。对于平面上的每一点 $Q=(x, y)$,存在唯一的对应点,即射线 NQ 与球面的第二个交点 P。反之,对应于球面上 N 点以外的每一点 P,存在具有有限坐标 (x, y) 的唯一点 Q。因此,可以把 P 看做为数 $x+\mathrm{i}y$ 的表示。当 P 以任何方式逼近北极 N 时,Q 就移向无穷;反之,如果 Q 以任何方式移向无穷,则对应的点 P 就逼近唯一确定的点 N。因而不对应于任何有限复数的这个点 N,自然要看做一切无穷大的 $x+\mathrm{i}y$ 的唯一表示,即看做为平面上的无穷远点,并附之以记号 ∞。如果不这样说,无穷远点就成为一个符号。引用此法,就把一切有限点和无穷远点在几何图形上完全等同了起来。

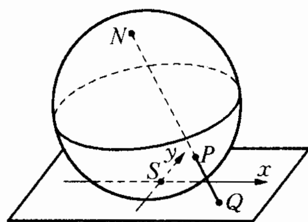


图 6.4

现在回过来谈代数关系(1)的几何解释。为此,也用一个 w 球面来代替 w 平面。这样,函数就可以像两个平面之间的映射一样,用一个从 z 球面到 w 球面的映射来表示。和两平面的映射的情况

一样,它也是保角的,因为根据一个众所周知的定理,平面到球面的立体射影是保角的。对应于 w 球面下唯一的一个位置,一般在 z 球面上有 n 个不同的位置。为了取得一一对应关系,再设想 w 球面上有 n 个叶,它们互相叠合,并通过适当的方法,借助于分支切割而组合成 w 球面上的一个 n 叶黎曼面。这个图并不比平面上的黎曼面更难懂。这样,代数方程(1)最终被解释成在 w 球面的黎曼面和 z 球面之间的一一对应关系,并且这种对应一般来说是保角的。这个解释显然也把 z 和 w 的无限值考虑在内了,它们可能彼此互相对应,或对应于有限值。

为了最充分地利用这个几何方法,必须采取一个相应的代数步骤,以排除无穷在公式中的例外作用,这个步骤就是引入齐次坐标,即令

$$z=\frac{z_1}{z_2},$$

并把 z_1, z_2 看做是独立的复变量,两者都是有限的,但不同时为零。于是,z 的每一个确定值就由无穷多组值 (cz_1, cz_2) 给出,其中 c 是一个任意常数因子。这样一组只差常数因子的值 (cz_1, cz_2),将被看做为在两个齐次变量域中的同一的位置。反之,对于每一个这样的位置,均有一个确定值 z,只有一个例外:对于(z_1 任意,而 $z_2=0$)这一位置没有有限的 z 与之对应,但如从其他位置向此位置逼近,则对应的 z 就变成无穷了。这个位置因而要看做是 z 平面上,也可以是 z 球面上,一个无穷远点的算术等价物,并记之为 $z=\infty$。

当然,同样地也置 $w=\frac{w_1}{w_2}$。现在要在齐次变量 z_1, z_2 和 w_1, w_2 之间建立齐次方程,以对应于方程(2)。此方程在乘以 z_2^n 去掉分母后可以写成

$$\frac{w_1}{w_2}=\frac{z_2^n\varphi\left(\frac{z_1}{z_2}\right)}{z_2^n\psi\left(\frac{z_1}{z_2}\right)}=\frac{\bar{\varphi}(z_1, z_2)}{\bar{\psi}(z_1, z_2)}。\tag{3}$$

在这个方程中，因为 $\varphi(z)$，$\psi(z)$ 最多含有 $z=\frac{z_1}{z_2}$ 的 n 次幂，所以 $\overline{\varphi}(z_1, z_2)$ 和 $\overline{\psi}(z_1, z_2)$ 是 z_1 和 z_2 的有理整函数。它们还是 n 次齐次多项式(形)，因为 $\varphi(z)$ 或 $\psi(z)$ 的每一项 z^i 去掉分母后都变成 n 次的项 $z_2^n\left(\frac{z_1}{z_2}\right)^i=z_2^{n-i}z_1^i$。

现在来详细研究由方程(1)(或方程(3))在 z 和 w 之间所建立的函数相依关系。我们将坚持使用两个新手段，即复球面上的映射和齐次坐标。如果对 z 球面和 w 球面上的黎曼曲面之间的保角关系作出一个完整的图来，这个问题就解决了。

首先必须问黎曼面分支点的性质和位置。我在这里提醒你们注意，一个 μ 重分支点就是 $\mu+1$ 个叶的连接点。由于 w 是 z 的单值函数，所以知道了对应于分支点的 z 球面点，也就知道了分支点。这种分支点，我习惯地称为 z 球面的临界点或值得注意的点。对应于这些点中的每一点，存在一个确定的重数，这个重数等于相应的分支点的重数。现在不加详细证明，只给出使这些点得以确定的定理。我假定你们一般都知道这里要用到的简单的函数论结果，尽管这些结果可能不是用我所喜欢的齐次形式来表述的。我将用具体的图形来阐明我要向你们介绍的抽象内容，在这方面举出一系列例子来。

必须稍微计算一下，以便求出在齐次坐标下与导数 $\frac{\mathrm{d}w}{\mathrm{d}z}$ 相当的量。对方程(3)微分，省略 $\overline{\varphi}$ 和 $\overline{\psi}$ 上的一横，得

$$\frac{w_2\,\mathrm{d}w_1-w_1\,\mathrm{d}w_2}{w_2^2}=\frac{\psi\mathrm{d}\varphi-\varphi\mathrm{d}\psi}{\psi^2}。\tag{3'}$$

我们还有

$$\mathrm{d}\varphi=\varphi_1\,\mathrm{d}z_1+\varphi_2\,\mathrm{d}z_2,$$

$$\mathrm{d}\psi=\psi_1\,\mathrm{d}z_1+\psi_2\,\mathrm{d}z_2,$$

其中

$$\varphi_1 = \frac{\partial\varphi(z_1, z_2)}{\partial z_1}, \quad \varphi_2 = \frac{\partial\varphi(z_1, z_2)}{\partial z_2},$$

$$\psi_1 = \frac{\partial\psi(z_1, z_2)}{\partial z_1}, \quad \psi_2 = \frac{\partial\psi(z_1, z_2)}{\partial z_2}。$$

另一方面，据 n 次齐次函数的欧拉定理，我们有

$$\varphi_1 z_1 + \varphi_2 z_2 = n\varphi,$$

$$\psi_1 z_1 + \psi_2 z_2 = n\psi。$$

因而(3′)式右侧的分子可记作形式

$$\psi \mathrm{d}\varphi - \varphi \mathrm{d}\psi = \begin{vmatrix} \mathrm{d}\varphi & \mathrm{d}\psi \\ \varphi & \psi \end{vmatrix}$$

$$= \frac{1}{n}\begin{vmatrix} \varphi_1 \mathrm{d}z_1 + \varphi_2 \mathrm{d}z_2 & \psi_1 \mathrm{d}z_1 + \psi_2 \mathrm{d}z_2 \\ \varphi_1 z_1 + \varphi_2 z_2 & \psi_1 z_1 + \psi_2 z_2 \end{vmatrix}。$$

根据行列式乘法定理，这个式子等于

$$= \frac{1}{n}\begin{vmatrix} \varphi_1 & \varphi_2 \\ \psi_1 & \psi_2 \end{vmatrix} \cdot \begin{vmatrix} \mathrm{d}z_1 & \mathrm{d}z_2 \\ z_1 & z_2 \end{vmatrix}。$$

这样，(3′)式转换为方程

$$\frac{w_2 \mathrm{d}w_1 - w_1 \mathrm{d}w_2}{w_2^2} = \frac{z_2 \mathrm{d}z_1 - z_1 \mathrm{d}z_2}{n\psi^2}(\varphi_1\psi_2 - \psi_1\varphi_2)。$$

这就是用来说明方程的齐次理论的基本公式，而 φ, ψ 的函数行列式 $\varphi_1\psi_2 - \varphi_2\psi_1$ 是随后一切推导的关键式子。除此式和因子 $z_2^2(n\varphi^2)$ 之外，右侧就是 $z = \frac{z_1}{z_2}$ 的微分，左侧是 $w = \frac{w_1}{w_2}$ 的微分。正如大家所知道的，对于有限的 z 和 w，临界点是由 $\frac{\mathrm{d}w}{\mathrm{d}z} = 0$ 给出的，所以下列定理是很令人信服的，但我们将略去其证明。函数行列式的每一个 μ 重零点是重数为 μ 的一个临界点，即与之相对应的是 w 球面上黎曼面的一个 μ

重分支点。这个规则与用其他方式给出的那些规则相比，其优点在于用同一个说法把 z 和 w 的有限值和无限值都包括进去了，它也对奇点的数目提出一个精确的说法。即是说，4 个导函数均为 $n-1$ 次的形式，因而函数行列式是 $2n-2$ 次的形式。如果把重数考虑在内，那么，这样的多项式就始终有 $2n-2$ 个零点。所以，如果 $\alpha_1, \alpha_2, \cdots, \alpha_v$ 是 z 球面的奇异点（即如果对于$z_1 : z_2 = \alpha_1, \alpha_2, \cdots, \alpha_v$，均有 $\varphi_1\psi_2 - \varphi_2\psi_1 = 0$），且 $\mu_1, \mu_2, \cdots, \mu_v$是它们相应的重数，则有

$$\mu_1 + \mu_2 + \cdots + \mu_v = 2n - 2。$$

借助保角映射，对应这些点，在 w 球面的黎曼面上存在着 v 个分支点

$$a_1, a_2, \cdots, a_v。$$

这些点必互相分离地位于黎曼面上，且分别有 $\mu_1 + 1, \mu_2 + 1, \cdots, \mu_v + 1$ 个叶相应围绕着它们循环连接。不过应该指出，这些分支点中的不同点可能处于球面的同一位置上，因为对于 $z = a_1, a_2, \cdots, a_v$ 的 $w = \dfrac{\varphi(z)}{\psi(z)}$，可能不止一次地给出相同的 w 值。在这样一个点上，可能有互相分离的两组以上的叶，每一组叶本身则是互相连接的。w 球面上的每一个这样的位置，称为分支位置，我们按顺序记之为 $A, B, C, \cdots$。应该指出，它们的数目可小于 v。

到此为止所作的说明，只能对黎曼面提供一个模糊的概念。现在要把它说完整，以便得到一个更直观的概念。为此，让我们在 w 球面上通过分支位置 $A, B, C, \cdots$画一条任意封闭曲线 l。这条曲线没有二重点，而且其形状要尽可能简单（图 6.5），并把由此形成的两个球冠分别称为上冠和下冠。在要讨论的一切例子中。点 $A, B, C, \cdots$都是实的，这样自然把实数的大子午圆选作曲线 l，所以上述两部分区域各成为一个半球。

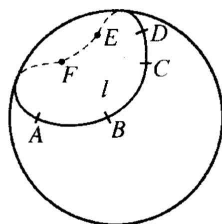


图 6.5

回到一般情况，我们看到，黎曼面上每一对连接的叶，沿着连接两个分支点的一个分支切口相交。你们知道，如果移动这些切口而保持其端点固定，即将这些叶看做为沿连接同样的分支点的其他曲线相连接时，黎曼面本质上仍然不变。黎曼面这个概念的巨大概括性，也是它很难掌握之处，就在于这种可变化性。为了给黎曼面以确定的形式，使人得到具体直观的形象，我们移动一切分支切口，使其全体处于通过一切分支点的曲线 l 上。可能有几个分支切口位于 l 的相同部分，而在 l 的另一些部分则可能根本没有任何切口。

现在沿着曲线 l 切割这整个叶丛，即把各个叶都切开。由于已经把所有的分支切口移到 l 上，因此刚才说的切割通过所有的分支切口，致使黎曼面分成了完全不再有分支的 $2n$ 个半叶，而在每个球冠上各有 n 个半叶。如果我们把对应上球冠的那些半叶设想成是加了阴影的，对应下球冠的是未加阴影的，从而简洁地区分为 n 个有阴影的半叶和 n 个无阴影的半叶。现在可以把原来的黎曼面描绘如下：在黎曼面上，每一个有阴影的半叶只是沿 l 的曲线段 AB，BC，…遇到与之连通的那些无阴影的半叶。同样地，每个无阴影的半叶仅仅沿着上述 l 的曲线段与有阴影的半叶连接。但两个以上的半叶只有可能在一个分支点相遇，而且事实上，围绕着任何 μ 重分支点，$\mu+1$ 个有阴影的半叶要同 $\mu+1$ 个无阴影的半叶交替地相遇。

由于 z 球面上的函数 $w(z)$ 对 w 球面上黎曼面的映射是一一对应的，因此马上把上述连通性条件转移到 z 球面上。由于存在着连续性，因此黎曼面的 $2n$ 个半叶必对应于 $2n$ 个连接的 z 的区域，这些区域分别称为有阴影和无阴影的半区域。这些区域沿 $z(w)$ 的 n 个像相互分离，因为 $z(w)$ 是 n 值函数，所以曲线 l 上的各段 AB，BC，…映射到 z 球面上各有 n 个像。每一个有阴影的半区域仅沿着这些像曲线相遇，而每一个无阴影区域只与有阴影的半区域相遇。两个以上的半区域只有在一个 μ 重的临界点上相遇。在这样一个点上，$\mu+1$ 个有阴影区和 $\mu+1$ 个无阴影区会合。

这种把 z 球面分割成 n 个局部区域的作法，有助于详细了解函数 $z(w)$ 的动向，我们将借以讲清 n 个简单的典型例子。下面我就从最简单的例子讲起。

1. “纯”方程

我们把众所周知的方程

$$z^n = w \tag{4}$$

称为纯方程，引入根号 $z=\sqrt[n]{w}$ 就得到了形式解。但是，这样完全没有讲清 z 和 w 之间的函数关系。因此，将根据一般方法引入齐次变量

$$\frac{w_1}{w_2}=\frac{z_1^n}{z_2^n},$$

考虑右侧分子和分母的函数行列式

$$\begin{vmatrix} nz_1^{n-1} & 0 \\ 0 & nz_2^{n-1} \end{vmatrix} = n^2 z_1^{n-1} \cdot z_2^{n-1}。$$

这个式子显然有 $n-1$ 重的零点 $z_1=0$ 及 $z_2=0$，它们在非齐次形式下就是 $z=0$ 及 $z=\infty$。这些是仅有的临界点，总重数为 $2n-2$。因此，根据一般定理，w 球面上黎曼面的仅有的分支点的位置为 $w=0$ 及 $w=\infty$，据方程 $w=z^n$，这些点对应于两个点 $z=0$ 及 $z=\infty$，每个点有重数 $n-1$，所以在各点上都有 n 叶交替地连接。现在在 w 球面上画出实数子午线作为曲线 l，并在一切分支切口适当移位后，把黎曼面的一切叶沿着这条子午线切开。在曲面所分成的 $2n$ 个半球面中，把 w 球面的后半部分，即对应于虚部为正的 w 的那些半球面，看做是有阴影的。子午线本身被区分为正实数半子午线(图 6.6 中实线部分)和负实数半子午线(图 6.6 中虚线部分)。

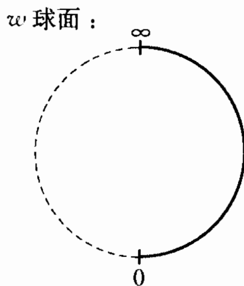


图 6.6

现在必须考察这条子午线即曲线 l 在 z 球面上的映像，z 球面就是用这些像典型地划分为半区域的。在正的半子午线上，$w=r$，r 取 0 到∞的正实数值。对于这些值，根据众所周知的复数公式，我们有

$$z=\sqrt[n]{w}=|\sqrt[n]{r}|\left(\cos\frac{2k\pi}{n}+\mathrm{i}\sin\frac{2k\pi}{n}\right),\ k=0,\ 1,\ \cdots,\ n-1。$$

对 k 的不同值，据此式可在 z 球面上得到 n 个半子午线，它们与正实数的半子午线的交角为 $0,\ \frac{2\pi}{n},\ \frac{4\pi}{n},\ \cdots,\ \frac{2(n-1)\pi}{n}$。因此，这些曲线对应于 l 中实线的那半部分。在 w 球面的负半子午线上，必须令 $w=-r=re^{i\pi}$，且仍有 $0\leqslant r\leqslant\infty$。由此得出

$$z=\sqrt[n]{w}=|\sqrt[n]{r}|\left(\cos\frac{(2k+1)\pi}{n}+\mathrm{i}\sin\frac{(2k+1)\pi}{n}\right)$$

$$k=0,\ 1,\ \cdots,\ n-1。$$

对应于此，在 z 球面上有 n 条半子午线，其“经度”分别为 $\frac{\pi}{n},\ \frac{3\pi}{n},\ \cdots,\ \frac{(2n-1)\pi}{n}$ ①，因而等分前述各子午线间的夹角。于是，z 球面从北极到南极被等分为 $2n$ 个全等的角域，类似于橘子的自然分瓣。这个切分与一般理论是一致的。特别是，两个以上的半平面区域，只是在奇异点即两极点上相遇。在上述每个点上，$2n$ 个半区域相遇，对应于重数 $n-1$。

为了给各区域打上阴影，只需把一个区域固定，其他部分交替地画上阴影或者不画。现在请注意，从点 $w=0$ 来看球面的有阴影的一半（即后部）时，实线分界部分在左侧，虚线部分在右侧。由于讲的是保角映像，所以角的定向不会反转，所以从 z 球面上的点 $z=0$ 来

---

① 英译本为 $\frac{2(n-1)\pi}{n}$ 有误。——编辑

看各阴影部分，位置关系也是相同的，即左侧有一实线的分界，右侧有一虚线的分界。这样就完全确定了 z 球面的区域划分。从对应于 $n=3$，$n=4$ 的情况，图 6.7 和图 6.8 中可以清楚地看到，两个 z 半球面上的区域分布因 n 为偶数或为奇数而有典型的区别。让我们强调一下，为了对这种情况获得一个全面的理解，很有必要转到复球面上去。在复的 z 平面上，将会得到由 $z=0$ 发出的射线作成的扇形划分，而完全看不清 $z=\infty$ 与 $w=\infty$ 和 $z=0$ 与 $w=0$ 分别作为临界点及分支点，具有相同的意义。

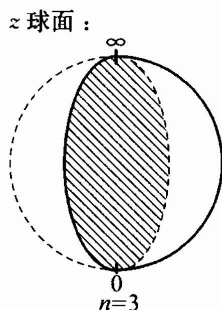


图 6.7

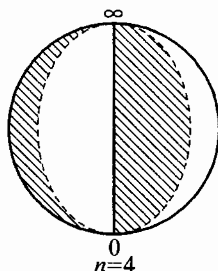


图 6.8

这就是为了准确了解 z 和 w 之间函数关系所需的要点。现在只要研究 $2n$ 个球面扇形中的各个球面扇形对两个 w 半球形中的某一个的保角映射。不过这里不准备详谈。这种情况是最简单和最明显的例子，是每个学过保角变换的人都熟悉的。在本章第五节中会看到，怎样用这个方法推导出 z 的数值计算。

然而这里要解决的是 z 球面上各个全等区域之间的相互关系这个重要问题。说得准确些，在 n 个有阴影的区域中，每一个区域中有一个点使 $w=z^n$ 取得同样的值。相应的 z 值，能否相互表示？事实上，我们注意到对于 $z'=z\cdot\varepsilon$（其中 ε 是 n 次单位根中的任何一个），$z'^n=z^n$，即 $w=z^n$ 在所有 n 个位置

$$z'=\varepsilon^{\nu}\cdot z=e^{\frac{2\nu i\pi}{n}}\cdot z\ (\nu=0,1,2,\cdots,n-1) \tag{5}$$

上取同样的值。z'的这 n 个值是这样分布的：如果 z 取自一个阴影区中的，则每个阴影区域中均各有一个 z'，而且当 z 移动穿过这个区域时，z'的 n 个值必然穿过各自对应的区域。无阴影区也是如此。代换(2)式中的每一个，用几何方法来表示，就是使 z 球围绕着垂直轴 0，∞旋转一个角$\frac{\nu \cdot 2\pi}{n}$，因为大家知道，复平面上乘以 $e^{\frac{2\nu i\pi}{n}}$，表示围绕原点旋转那样一个角度。因此，围绕垂直轴作这 n 个的旋转，球面区域中相应的点以及区域本身就会发生相互转换。

如果说我们一开始只确定球面上一个有阴影部分区域，那么按上面的说明法，可以得出一切类似的部分球面区域。这里只利用了代换(5)的这样一个性质，即它使方程(4)变为自身(即 $z^n = w$ 转换成 $z'^n = w$)，且代换的数目等于方程的次数。在下面的例子中，就总是可以一开始就给出那样的线性代换，使子区域的分割得以简化。

我要用这个例子来说明一个重要的一般概念，即有理地含有参数 w 的方程的不可约性。我们已经结合正七边形的构造讨论过带有理系数的方程的不可约性(参看第一部分第三章)。设有一个方程 $f(z, w) = 0$ (即我们的方程 $z^n - w = 0$)，其中 $f(z, w)$是关于 z 的一个多项式，其系数为 w 的有理函数，则当 f 可分解为两个同类的多项式的乘积：

$$f(z, w) = f_1(z, w) \cdot f_2(z, w)$$

并且右侧每一多项式都真正含有 z 时，称此方程就参数 w 而言是可约的，否则称就参数 w 而言不可约。与以前的概念相比，这个概念的全部推广就在于：我们进行运算的，同时也是上述关于 z 的多项式系数所从属的“有理性域”，是由参数 w 的全部有理函数构成，而不是由全部有理数构成。换句话说，我们从数论的概念转移到了函数论的概念。

如果对于每个方程 $f(z, w) = 0$ 都借助于黎曼面来说明这一点，我们就可以定出在新意义下可约性的一个简单标准。例如，如此

方程可约，则满足此方程的每一组值 z, w，不是满足 $f_1(z, w)=0$，就是满足 $f_2(z, w)=0$。现在，$f_1=0$ 和 $f_2=0$ 的解都可用黎曼面来表示，它们彼此毫无关系，特别是互不连通。这样，属于可约方程 $f(z, w)=0$ 的黎曼面，至少必分成两个分离的部分。

据此可以断言，方程 $z^n-w=0$ 在函数论的意义下肯定是不可约的，因为已确知在黎曼面的每一个分支点上，n 个叶都是循环相通的。此外，整个黎曼面映射在未分割的 z 球面上，因此，这种分离是不可能发生的。

与此相关，可以回答前面涉及的一个流行的数学问题，即任意角 n 等分的可能性问题，特别是对 $n=3$，即可否将角三等分的问题。这个问题就是要用圆规和直尺精确作图，把任意角 φ 三等分（当然很容易对 φ 的一系列特殊值作图）。我将提供一个思路，证明不可能按上述的意义将角三等分。为此请回忆正七边形不可能用圆规直尺作图的证明（参看第一部分第三章）。和在那里完全一样，将把问题归结为解不可约三次方程问题，然后证明这个方程不能用一系列平方根来解。不同之处在于，现在的方程含有一个参数（角 φ），而以前的方程的系数却是整数。因此，必须用函数论上的不可约性取代数论上的不可约性。

为了建立这个问题的方程，请把角 φ 看成是从 w 平面上正实半轴算起的（图 6.9），因而它的动边在点

$$w=e^{i\varphi}=\cos\varphi+i\sin\varphi$$

处与单位圆相交。我们的问题是：对参数 φ 的任意值，找出一个使用有限次圆规、直尺的作图法，求出角 $\frac{\varphi}{3}$ 的边与单位圆的相交点，即求出点

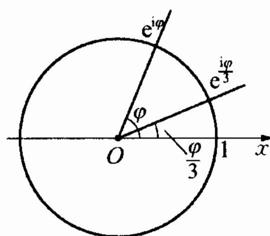


图 6.9

$$z=e^{\frac{i\varphi}{3}}=\cos\frac{\varphi}{3}+i\sin\frac{\varphi}{3}。$$

z 的这个值满足方程

$$z^3 = \cos\varphi + \mathrm{i}\sin\varphi, \tag{6}$$

而与这个几何问题等价的代数问题，是用 $\sin\varphi$ 和 $\cos\varphi$ 的有理函数的有限次相互叠置的平方根，解出这个方程（见第二部分第三章）。因为这些量是开始作图时的点 w 的坐标。

必须首先证明方程(6)在函数论的意义上是不可约的。肯定地，这个方程并不具有在解释这个一般概念时所假定的形式，因为其中不是有理地含有一个复参数 w，而是含有一个实参数 φ 的两个函数 $\cos\varphi$ 和 $\sin\varphi$，而且是有理地含有这两个函数。作为上述概念在这里的自然延伸，如果多项式 $z^3-(\cos\varphi+\mathrm{i}\sin\varphi)$ 能分解成若干个系数为 $\cos\varphi$ 和 $\sin\varphi$ 的有理函数的多项式，即可称其为可约。同时也能像以前一样，为此建立一个准则。如果在(6)式中让 φ 取遍所有实数值，那么，$w=\mathrm{e}^{\mathrm{i}\varphi}=\cos\varphi+\mathrm{i}\sin\varphi$ 的图形就是 w 平面上的单位圆，通过球极射影变换，在 w 球面上有一方程对应于它。此单位圆在方程 $z^3=w$ 的黎曼曲面上所对应的曲线是可以用一笔划出的三叶，而由(6)式唯一地映到 z 球面的单位圆上。因此，在某种意义上，它可以当作“一维黎曼映像”来看。用同样方法，对每一个形如 $f(z, \cos\varphi, \sin\varphi)=0$ 的方程，可按如下方式构造这样一种黎曼映像与之对应：方程有多少个根，就取多少个具有弧长为 φ 的单位圆并按各个根的连接情况把它们连接起来。正如前面一样，由此可以推出，仅当方程(6)的一维黎曼映像能分离成几个部分时，该方程才是可约的，而现在显然不是这种情况。这就证明，方程(6)在函数论的意义上不可约。

不过，以前定理的证明，即有理系数的三次方程，当它可用一系列平方根求解时方为可约，完全可以逐字地用到函数论意义下不可约的方程(6)上来。只需要将该处的“有理数”代之以“$\cos\varphi$ 与 $\sin\varphi$ 的有理函数”即可。这就证明了我们关于“一个任意角不能用有限次圆规、直尺作图来三等分”的论断。因此，三等分角的狂热者必徒劳无功！

现在转到处理一个较为复杂的例子。

2. 二面体方程

所谓二面体方程就是方程

$$w=\frac{1}{2}\left(z^{n}+\frac{1}{z^{n}}\right), \tag{7}$$

理由容后述。消去分式即可看出其次数为 $2n$。引入齐次变量,得

$$\frac{w_1}{w_2}=\frac{z_1^{2n}+z_2^{2n}}{2z_1^{n}\cdot z_2^{n}},$$

其分子和分母都是 $2n$ 次形式。这些形式的函数行列式是

$$\begin{vmatrix} 2nz_1^{2n-1} & 2nz_2^{2n-1} \\ 2nz_1^{n-1}z_2^{n} & 2nz_1^{n}z_2^{n-1} \end{vmatrix}=4n^2z_1^{n-1}z_2^{n-1}(z_1^{2n}-z_2^{2n})。$$

这个行列式在 $z_1=0$ 和 $z_2=0$ 上具有 $n-1$ 重零点,其他 $2n$ 个零点由下式给出

$$z_1^{2n}-z_2^{2n}=0\quad 或\left(\frac{z_1}{z_2}\right)^{n}=\pm 1。$$

如果除了已经用过的 1 的 n 次根

$$\varepsilon=e^{\frac{2i\pi}{n}}$$

以外,再引入 -1 的 n 次元根

$$\varepsilon'=e^{\frac{i\pi}{n}},$$

则这 $2n$ 个零点由下式给出

$$\frac{z_1}{z_2}=\varepsilon^{\nu},\ \frac{z_1}{z_2}=\varepsilon'\varepsilon^{\nu}(\nu=0,\ 1,\ \cdots,\ n-1)。$$

既然对应于这些点的值 $z=\frac{z_1}{z_2}$,其绝对值都等于 1,可知它们都位于

z 球面的赤道上(对应于 z 平面的单位圆),其角度间隔均为$\frac{\pi}{n}$。因而作为 z 球面的临界点,我们有

(a) 南极 $z=0$ 和北极 $z=\infty$,重数各为 $n-1$。

(b) 赤道上的 $2n$ 个点 $z=\epsilon^{\nu}$, $z=\epsilon'\epsilon^{\nu}$,重数各为 1。

重数之和为 $2(n-1)+2n\cdot 1=4n-2$,符合前面关于 $2n$ 次的一般定理的要求。借助于方程(7),对应于 z 球面上的特殊的点 $z=0$, $z=\infty$,都有 w 球面上的$w=\infty$。此外,对应于一切点 $z=\epsilon^{\nu}$,有 $w=+1$;对应于一切点 $z=\epsilon'\cdot\epsilon^{\nu}$,有 $w=-1$。因而,在 w 球面上只有 3 个分支点∞, $+1$, -1。这些点的位置如下:

$w=\infty$,两个重数为 $n-1$ 的分支点;

$w=+1$, n 个重数为 1 的分支点;

$w=-1$, n 个重数为 1 的分支点。

因此,黎曼面有 $2n$ 个叶,其组合情况如下:在点 $w=\infty$ 处,分为两组,每组 n 个叶,循环相连;在 $w=+1$ 和 $w=-1$ 处,各分为 n 组,每组两叶。研究了 z 球面相应地分割成半区域后,就会明白这些叶的排列。

为此目的,正如前面所说的,最好了解一下将方程(7)变成自身的线性变换。像纯方程的情况一样,方程(7)在下述 n 个代换

$$z'=\epsilon^{\nu}z\ (\nu=0,1,\cdots,n-1),\text{其中}\ \epsilon=e^{\frac{2i\pi}{n}} \tag{8a}$$

下不变,因为对这些代换, $z'^n=z^n$。同样,在另外 n 个代换

$$z'=\frac{\epsilon^{\nu}}{z}\ (\nu=0,1,\cdots,n-1) \tag{8b}$$

下,它也不变,因为这些代换只是把 z^n变为$\frac{1}{z^n}$。

因此,我们有 $2n$ 个线性代换使方程(7)变为自身,总个数恰好和方程(7)的次数相等。如果对于 w 的一个给定值 w_0 知道方程的一

个根 z_0，立刻就知道 $2n$ 个根 $\varepsilon^\nu \cdot z_0$ 和 $\frac{\varepsilon^\nu}{z_0}$ $(\nu = 0, 1, \cdots, n-1)$。虽然它们与 w 的同一个值 w_0 相对应，但这些根一般是不同的。也就是说，求出了 n 次单位根 ε 以后，也就知道了方程的全部根。

现在来研究若将 w 球面的黎曼面沿实子午线切割，z 球面怎样分割。就像前一个例子一样，这里将 w 球面的实子午线按各分支点分为 3 段，一段是由 $+1$ 到 ∞（实线标出部分），另一段是由 ∞ 到 -1（细点线标出部分），还有一段是 -1 到 $+1$（短线段标出部分）（图6.10）。与这 3 段中的每一段相对应的，有 z 球面上的 $2n$ 个不同的曲线段，它们可以通过 $2n$ 个线性变换(8)从其中任一段导出。因此，只要求这些线段中的一个就足够了。此外，所有这些线段必然连接临界点 $z = 0, \infty, \varepsilon^\nu, \varepsilon' \cdot \varepsilon^\nu$，所以我们在 z 球面上标明它们。正像前一个例子一样，因 n 的奇偶不同，其形式也略有不同。这只要拿一个确定的情况例如 $n = 6$ 来说明就够了。图 6.10 是正交射影中的 z 球面前半部分。我们看到，在赤道上，从左到右相隔 60°，分别有 $\varepsilon^3 = -1$, ε^4, ε^5, $\varepsilon^6 = 1$，在它们的中间等分位置上还有 $\varepsilon' \cdot \varepsilon^3$, $\varepsilon' \cdot \varepsilon^4 = i$ 和 $\varepsilon' \cdot \varepsilon^5$。

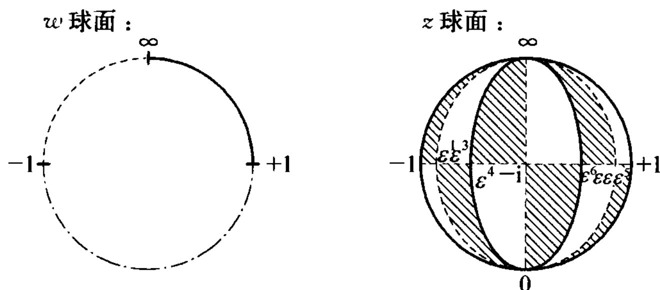


图 6.10

现在我们将看到，实 z 子午线的象限圆 $+1 < z < \infty$，对应于实 w 子午线的 $+1 < w < \infty$ 部分（实线标明的部分）。事实上，如果置 $z = r$，令 r 从 1 到 ∞ 取实数值，那么 $w = \frac{1}{2}\left(z^n + \frac{1}{z^n}\right) = \frac{1}{2}\left(r^n + \frac{1}{r^n}\right)$

也将从 1 取实数值增加到∞，从这条曲线，通过 n 个线性变换(8a)，得出 z 球面上由实线标明的其他几条曲线。但是，正如在前一例中所看到的，这些代换意味着球面围绕着垂直轴$(0,\ \infty)$分别转过角$\frac{2\pi}{n}$，$\frac{4\pi}{n}$，…，$\frac{2(n-1)\pi}{n}$。这样从北极∞到赤道上的点 ε^{ν}，得到 n 个子午线象限圆。如果运用代换 $z'=\frac{1}{z}$，把子午线上从$+1$ 到∞的象限圆变换成从$+1$ 到 0 的实子午线下部的象限圆，就得到另一条用实线标明的曲线。如果对这个象限圆进行 n 个旋转(8a)，加上 $z'=\frac{1}{z}$，合成 n 个代换(8b)，就得出另外 n 个子午线象限圆，它们把南极同赤道点 ε^{ν} 连接起来。现在事实上有了 $2n$ 条以实线标明的曲线，对应于 w 球面上以实线标明的象限圆。对于 $n=6$ 的特殊情况，这些曲线组成 3 条完整的子午线，由实子午线经旋转 0°，60°，120°而成。

现在很明显，当 r 取遍从$+1$ 到∞的实数值时，值 $z=\varepsilon'\cdot r$ 的全体对应于实子午线 w 上以点线标明的部分，因为此时由方程(7)得出

$$w=\frac{1}{2}\left(\varepsilon'^{n}r^{n}+\frac{1}{\varepsilon'^{n}r^{n}}\right)=-\left(r^{n}+\frac{1}{r^{n}}\right)。$$

这个式子实际上是通过实数值从-1 下降到$-\infty$的，但 $z=\varepsilon'^{r}$ 所表示的是从∞到赤道上的点 ε^{ν} 的子午线象限圆。如果现在对它实施代换(8a)、(8b)，像前面一样，就会发现，对应于实子午线 w 的以点线标明的部分，将是把南北极同赤道点 $\varepsilon'\cdot\varepsilon^{\nu}$ 连接起来的所有子午线的象限圆，它们将前面得出的那些子午线象限圆之间的角等分。对于 $n=6$ 的特殊情况，它们由实子午线经过 30°，90°，150°的旋转所变成的 3 条完整的子午线组成。

剩下来的问题就是求对应于以短线段标明的半子午线 $-1<$

$w<+1$ 的 $2n$ 个曲线段。我们将证明，它们是 z 球面赤道上连接点 ε^{ν} 和 $\varepsilon' \cdot \varepsilon^{\nu}$ 的曲线段。事实上，赤道所表示的是绝对值为 1 的点，因而由 $z=e^{i\varphi}$ 给出，其中 φ 是从 0 至 2π 的实数。由此得

$$w=\frac{1}{2}\left(z^{n}+\frac{1}{z^{n}}\right)=\frac{1}{2}\left(e^{ni\varphi}+e^{-ni\varphi}\right)=\cos n\varphi。$$

这个表示式始终是实的，其绝对值不大于 1。事实上，当 φ 由 $\frac{\pi}{n}$ 的一个倍数变到下一个倍数，即 z 走过这里所讲的曲线段之一时，它取 $+1$ 与 -1 之间的每一个值一次。

以这种方式确定的曲线，把 z 球面分成 $2\cdot 2n$ 个三角半区域，并以 3 条曲线各一条为界，而各半区域又对应于黎曼面的半叶。几个区域只能在临界点上相遇。于是根据本节前面提到的重数表，北极和南极为 $2n$ 重，每个点 ε^{ν} 和 $\varepsilon'\varepsilon^{\nu}$ 为 $2\cdot 2$ 重。为了确定这些区域中哪一些是要加阴影的区域，我们注意到，当 w 按实线、短线段线和点线标明的实子午线顺序移动时，w 球面的后一半处在它的左边。由于映射是保角的，因此应加阴影的是那些边界按上述走向的那些半区域，而其余部分则不加阴影。

现在已对方程中所建立的 z 和 w 的相互依赖关系得出一个完整的几何图形。本来可以更详尽地加以解释，进一步探讨将单个三角形域映射成 w 半球的保角变换关系，但不这样做了，我只扼要描述一下 $n=6$ 的情况，对这个情况我已经给了特别的注意。此时，z 球面可以分为 12 个有阴影的三角和 12 个无阴影的三角，其中每一种三角有 6 个可以在图 6.11 中看到。每一种三角中有 6 个在一个极上相遇，有两个在赤道的 12 个等距离点中的一点上相遇。每个三角形被保角映射到同类的一个 w 半叶上。在黎曼面的半叶中，每一类有 6 个在分支位置∞上连接，有两个在每一个分支位置± 1 上连接，这对应 z 球面上半区域的分组。

我们可以得到关于 z 球面划分的一种方便图形，由于它同即将

出现的一些图形有类似之处,因此特别有价值。现在说明如下:如果把赤道上 n 个等距离点(即 ε^{ν})用直线按顺序连接起来,并把其中的每一个点同两极连接起来,那么就得到了一个对顶棱锥,它有 $2n$ 个面且内接于球面(图 6.11 中有 12 个面);如果现在从中心把 z 球面的区块投影到对顶棱锥上,就可以把每一个锥面用从极点作的高分成有阴影和无阴影的两半。如果用这个对顶棱锥来表示 z 球面的划分,因而用以表示我们的函数,那么对顶棱锥的作用就同下面要出现的正多面体的例子中的作用相类似。如果设想把对顶棱锥压到它的底面上去,并考虑由此产生的双重正 n 边形(六边形),其上下两面被通过其中心和顶点及各边中点的直线各自分为 $2n$ 个三角形(图 6.12),那么情形就完全相同了。我习惯于称这个图形为二面体,并将它看做与柏拉图时代以来一直在研究的 5 个正多面体属于一类。它事实上满足定义正多面体的一切条件,因为它的面(n 边形的两面)是全等的正多边形,而且具有全等的边(n 边形的边)和全等的顶点(n 边形的顶点)。唯一区别是它不是真正立体的边界,它包含的体积为 0。因此,柏拉图定理之所谓仅有 5 个正立体,只有在定义中加进指出真正立体的条件下才是正确的,而在证明中始终隐含着这个要求。

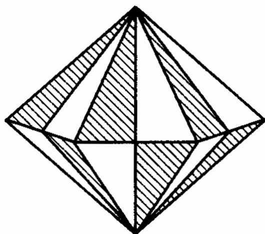


图 6.11

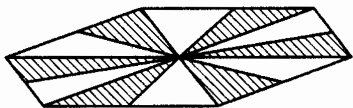


图 6.12

如果从二面体出发,将之投影到该球面上,不仅把它的顶点,而且把它的边和面的中心都投影到球面上,但后者的投影射线要垂直于二面体的平面,那么就得到 z 球面的划分。因此也可以认为这个

二面体表示按我们的方程所建立的在 w 和 z 之间的函数关系。由此可见,已经用过的那个简单扼要的名称——二面体方程,是恰当的。

下面就要考虑已经提到过的、与柏拉图正多面体密切相关的那些方程。

3. 四面体、八面体、二十面体方程

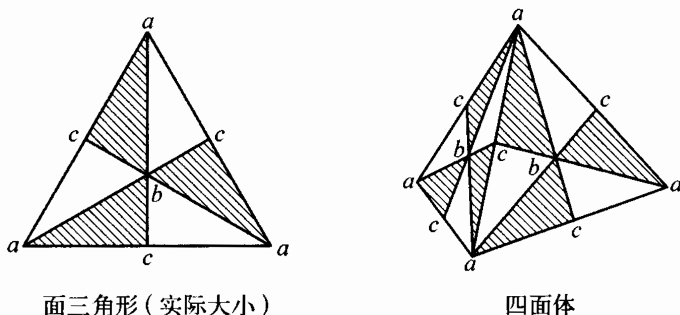
我们将看到,后两个方程同样可称为六面体和十二面体方程。这样的话,所有 5 种正多面体都包括进去了。在这一节里,将走同前一节例子相反的路子。我们将从正多面体出发,导出球面的区域划分,然后建立适当的代数方程,使球面分区图形成为对代数方程的恰当的几何解释。但我时常不得不限于做一些提示,遇到这种地方,请立刻参考我的书:*Vorlesungen über das Ikosaeder und die Auflösung der Cleichungen vom fünften Grade*①。在那本书里,可以找到对范围很广的整个理论的系统说明,以及它与有关领域的关系。

此外,将平行地讲解 3 种情况,并先从四面体的球面剖分开始。

(1) 四面体(图 6.13)。把四面体的 4 个等边三角形各用其 3 个高各分成 6 个小三角形,这 6 个小三角形可分成两组,每组有 3 个彼此全等的三角形,而任两个相互不全等的三角形是对称的。这样就把整个四面体的面分成了 24 个三角形,属于两组,各含有 12 个全等的三角形,而一组内的任一个对称于另一组的任一个。我们把一组的三角形画上阴影。这 24 个三角形的顶点可分成下述 3 类,使每个三角形在每一类顶点中都具有一个顶点:

---

① 莱比锡,1884 年。后面提到时,称《二十面体》。此书曾由 G · C · 莫里斯(G. C. Morrice)译成英文,名为:克莱因著《二十面体讲义》,1911 年修订版,基根 · 保罗(Kegan Paul)股份公司。



面三角形(实际大小)　　　　四面体

图 6.13

(a) 四面体的原来 4 个顶点,在每一个顶点处均有 3 个有阴影和 3 个无阴影的三角形相遇。

(b) 4 个面的 4 个重心,它们决定了另一个正四面体(所谓共四面体);在这 4 个重心中的任一个重心上,每一组均有 3 个三角形相遇。

(c) 棱的 6 个中点,决定了一个正八面体;在这 6 个中点的每一点,每一组中各有两个三角形相遇。

如果从四面体的重心出发,将此三角形分割投影于外接球面上,那么外接球面又可分为 2 · 12 个三角形,这些三角形都以大圆弧为边界,或互为全等或彼此对称。环绕着(a)、(b)、(c)类的每一个顶点,相应地有 6 个、6 个、4 个相等的角。因为环绕球面上一个点的角度之和为 2π,所以每个球面三角形在(a)或(b)类顶点的顶角为$\frac{\pi}{3}$,而在(c)类顶点的顶角为$\frac{\pi}{2}$。

这种球面划分的典型特征,还有四面体本身的特征,都是围绕其中心作一系列旋转可将它们变换成自身。仔细看一下那个四面体及其划分的模型就能明白。对本章而言,只要指出可能的旋转的个数就足够了(这里,静止位置也作为恒等旋转而被包括在内)。如果取原四面体的一个确定的顶点,那么,通过旋转,此顶点就可以变换为

四面体的任一个顶点(包括它本身),一共有 3 个可能性[①]。若将此顶点固定在这 4 个位置中的任一个位置上,则仍能使四面体旋转变换。因此共有 4 · 3=12 个旋转,把四面体或外接球面的相应三角形划分转换成其自身。通过这些旋转,可以使预先指定的有阴影(或无阴影的)三角形转换成其他任一个有(或无)阴影的三角形。取定第二个三角形之后,特定的旋转也就确定了。这 12 个旋转显然形成被称为G_{12}的群,其中含有 12 个运算。所谓群即是说,若逐次进行其中两个运算,其结果为 12 个旋转之一。

如果把这个球面看做 z 球面,那么这 12 个旋转的每一个都能以 z 的线性变换来表示,这样产生的 12 个线性变换就会使对应于四面体的方程变换为自身。为了比较起见,我要指出,也可把二面体方程的 $2n$ 个线性变换解释为使二面体旋转为自身的旋转的总体。

(2) 现在可以用类似的方式处理八面体(图 6.14),这样也许简单一些。可以像前面一样,把其中每个面分成 6 个小三角形,由此得到这样一种划分:把八面体整个表面分成全等的 24 个有阴影的和 24 个无阴影的全等三角形,而无阴影的三角形与有阴影的三角形互相对称。其 3 类顶点为:

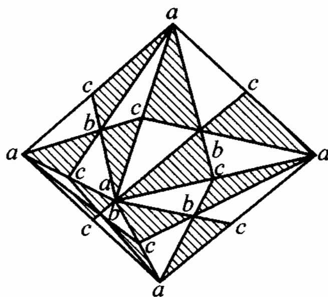


图 6.14

(a) 八面体的 6 个顶点,每类三角形中各有 4 个在这些顶点相遇。

(b) 8 个面的 8 个重心,它们形成一个立方体的顶点,在这些顶点上两类三角形每类各有 3 个相遇。

(c) 各边的 12 个中点,在每个中点,两类三角形每类各有两个

① 英文版原文作"四",不对,从下面的计算(4 · 3=12)也可看出。——中译者

相遇。

如果现在来研究外接球面，那么通过中心射影，可以得到 2 · 24 个球面三角形的划分，这些三角形或为全等或为对称，其中每一个三角形在(a)类顶点处的角是$\frac{\pi}{4}$，在(b)类顶点处的角是$\frac{\pi}{3}$，在(c)类顶点处的角是$\frac{\pi}{2}$。由于顶点 b 形成一个立方体，所以很容易看出，如果从立方体出发，把顶点和它的各面和各边的中心投影到球面上，那么在球面上也可以得到同样的划分。换句话说，我们不必再对这个立方体给予特别的注意。

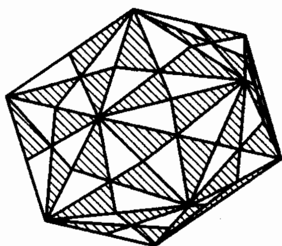


图 6.15

正如前面一样，很容易看出，通过形成群 G_{24} 的 24 个旋转变换，八面体以及球面的划分将变换成自身。而且，一旦预先指定的有阴影三角形被转换为另一确定的有阴影三角形时，这个旋转就完全确定下来了。

(3) 最后考虑正二十面体(图 6.15)。也同样从 20 个三角形的面①的划分开始，得到 60 个有阴影的和 60 个没有阴影的小三角形。它们具有如下 3 类顶点：

(a) 二十面体的 12 个顶点，每类三角形各有 5 个在此点相遇。

(b) 二十个面的 20 个中心，它们组成正十二面体的顶点，每类三角形各有 3 个在这类点相遇。

(c) 所有边的中点共 30 个，每类三角形各有两个在这类顶点相遇。

过渡到球面上时，每个球面三角形在(a)类、(b)类、(c)类顶点相

① 英文版原文误为“24 个三角形面”。——中译者

应的顶角是$\frac{\pi}{5}$，$\frac{\pi}{3}$，$\frac{\pi}{2}$。从(b)类顶点的性质，可如前面的考虑一样，得知考虑正十二面体，也可以得到同样的划分。

最后，二十面体及其对应的球面划分通过围绕其中心的60个旋转所成的变换群G_{60}而变换成自己。只要看一个模型，你们就会明白这些旋转，同时也会明白正八面体的旋转。

现列出已考虑过的3种情况下球面三角形的角(这里再加上二面体)：

二面体：$\frac{\pi}{2}$，$\frac{\pi}{2}$，$\frac{\pi}{n}$；

四面体：$\frac{\pi}{3}$，$\frac{\pi}{3}$，$\frac{\pi}{2}$；

八面体：$\frac{\pi}{4}$，$\frac{\pi}{3}$，$\frac{\pi}{2}$；

二十面体：$\frac{\pi}{5}$，$\frac{\pi}{3}$，$\frac{\pi}{2}$。

我可以把库默尔说过的一个笑话变一个样子说，搞自然科学的人会立刻由此得出结论，以为球面还可以作其他的划分，并使划分的结果具有类似的性质，而具有$\frac{\pi}{6}$，$\frac{\pi}{3}$，$\frac{\pi}{2}$，$\frac{\pi}{7}$，$\frac{\pi}{3}$，$\frac{\pi}{2}$之类的顶角。当然数学家是不会冒类此推断的危险的，他在这种地方小心从事是有道理的，因为这类球面划分事实上已列举殆尽，再不可能有其他分法了。这与以下的事实当然有关系：不存在其他的正多面体。我们可以在整数的一个性质中看出这个事实的终极解释，而不可再归结为更简单的解释。这个解释就是球面三角形的每个角必须是π的整数等分中的一份，例如说$\frac{\pi}{m}$，$\frac{\pi}{n}$，$\frac{\pi}{r}$，且分母要满足不等式

$$\frac{1}{m}+\frac{1}{n}+\frac{1}{r}>1。$$

这个不等式有一个性质，即它的整数解只能是上表列出的那些。这

个不等式很容易理解,因为它只不过表明了一个事实,即球面三角形各角之和大于 π。

我想在这里提一下,正如你们中某些人无疑已知道的,这个理论的适当推广也不会越过这些看来太狭窄的范围很远:自守函数理论涉及将球面划分成无穷多个三角形,而其内角和小于或等于 π。

4. 续:建立正规方程

现在进入问题的第二部分,建立形式如下方程

$$\varphi(z)-w\psi(z)=0 \quad 或 \quad w=\frac{\varphi(z)}{\psi(z)}。 \tag{9}$$

这个方程相应属于球面的上述 3 种划分,即将 w 球面的两个半球射影为 z 球面上 2 · 12,2 · 24 或 2 · 60 个小三角形。对每个 w 值,一般必有对应的 12,24,60 个 z 值,每一个值都处于适当类型的一个三角形之内。因此,所求的方程在这 3 种情况下,其次数必然分别为 12,24,60,我们将一般地记为 N。现在,每个局部区域接触 3 个临界点;因此,在每种情况下,w 球面上必然都有 3 个支点。按习惯指定其为 $w=0,1,\infty$,且再度选取通过这 3 点的实子午线为切口曲线 l,它的 3 段将对应于 z 球面上三角形的边界。

在 3 种情况的任一情况下,都取各面的重心(按以前的记号是(b)类顶点)对应于点 $w=0$,各边中点((c)类顶点)对应于点 $w=1$,多面体的顶点((a)类顶点)对应于点 $w=\infty$(图 6.16)。于是三角形的各边按指定的方法被映射成 w 子午线的 3 个线段,有阴影的三角形对应于 w 的后半球,无阴影的三角形对应于 w 的前半球。根据这些对应关系,方程(9)就是将 z 球面映像成以 0,1,∞为支点的 w 球面的 N 叶黎曼面的单值映射。

可以借助一般函数论的理论先验地推导出这个方程存在性的证明,但我宁可不假设读者具有这个方法所需的知识,而用经验方法去构造各种方程。这个方法也许能使我们对各个情况形成更生动的直观。

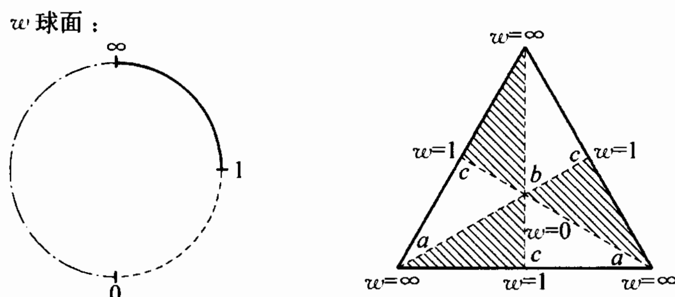


图 6.16

将方程(9)写成齐次变量形式：

$$\frac{w_1}{w_2}=\frac{\Phi_N(z_1,\ z_2)}{\Psi_N(z_1,\ z_2)},$$

其中 Φ_N，Ψ_N 是 z_1，z_2 的 N ($N=12$，24 或 60) 次齐次多项式。在这种形式的方程中，w 球面上的点 $w_1=0$，$w_2=0$(即 $w=0$，∞) 似乎比第三个支点 $w=1$ (齐次形式为 $w_1-w_2=0$) 更受重视。但是就我们的目的而言，由于 3 个支点是同等重要的，不妨也考虑下面形式的方程

$$\frac{w_1-w_2}{w_2}=\frac{X_N(z_1,\ z_2)}{\Psi_N(z_1,\ z_2)},$$

其中分子 $X_N=\Phi_N-\Psi_N$ 也是 N 次形式。这两个方程可以都包括在连比例式

$$\begin{aligned}&w_1:(w_1-w_2):w_2\\&=\Phi_N(z_1,\ z_2):X_N(z_1,\ z_2):\Psi_N(z_1,\ z_2)\qquad(10)\end{aligned}$$

中，这就为方程(9)提供了完整的齐次式，对所有的支点作出了同等的考虑。

现在的问题是建立函数 Φ_N，X_N，Ψ_N。为此目的，将它们和 z 球面上的有关划分联系起来。从方程(10)可看到，$\Phi_N(z_1,\ z_2)=0$

对应于 $w_1=0$,即 $w=0$ 对应于 Φ_N 在 z 球面上的 N 个零点。另一方面,多面体各面的重心(划分的(b)类顶点)在各种情况下都有 $\frac{N}{3}$ 个,据我们的假设,它们必须对应于支点 $w=0$。每一个重心必然是我们方程的三重根,因为 z 球面上总有 3 个有阴影和 3 个没有阴影的三角形在每个这样的点相遇。因此,这些重数为 3 的点,提供了所有对应于 $w=0$ 的位置,亦即 Φ_N 的所有零点。Φ_N 只含三重零点,故必为某 $\frac{N}{3}$ 次函数 $\varphi_N(z_1, z_2)$ 的三次方

$$\Phi_N=[\varphi_{\frac{N}{3}}(z_1, z_2)]^3$$

用同样的方法可推出,$X_N=0$ 的零点对应于 $w=1$(即 $w_1-w_2=0$)的位置,且等同于多面体各边的 $\frac{N}{2}$ 个中点(据我们划分,为(c)类顶点),每个中心计算两次。因而,X_N 必然是某个 $\frac{N}{2}$ 次函数的平方

$$X_N=[\chi_{\frac{N}{2}}(z_1, z_2)]^2。$$

最后,Ψ_N 的零点对应于点 $w=\infty$,因而必然要等同于多面体的顶点((a)类顶点),但在这些顶点上,因各种情况而分别有 3 个、4 个或 5 个三角形相遇,因而我们有

$$\Psi_N=[\psi_{\frac{N}{v}}(z_1, z_2)]^\nu, \quad 其中 \nu=3, 4 或 5。$$

我们的方程(10)因而必然具有形式

$$\begin{aligned} & w_1:(w_1-w_2):w_2 \\ = & \varphi(z_1, z_2)^3:\chi(z_1, z_2)^2:\psi(z_1, z_2)^\nu, \end{aligned} \tag{11}$$

其中 φ, χ, ψ 的次数与幂数及方程的次数 N 列于下面:

四面体:$\varphi_4^3, \chi_6^2, \psi_4^3$;$N=12$;

八面体:$\varphi_8^3, \chi_{12}^2, \psi_6^4$;$N=24$;

二十面体:$\varphi_{20}^3, \chi_{30}^2, \psi_{12}^5$;$N=60$。

我现在要简略地证明,已讨论过的二面体方程也能归入格式(11)之内。只要回忆一下,在此情况下,我们选择−1, +1, ∞,而不是后来取的 0, +1, ∞,作为 w 球面上的支点。这样,只要把二面体方程改为

$$(w_1+w_2):(w_1-w_2):w_2=\Phi:X:\Psi,$$

就可以得到与(11)式完全类似的结果。现在从用过的二面体方程

$$\frac{w_1}{w_2}=\frac{z_1^{2n}+z_2^{2n}}{2\cdot z_1^n\cdot z_2^n}$$

出发,通过化简,可得

$$\begin{aligned}&(w_1+w_2):(w_1-w_2):w_2\\=&(z_1^{2n}+z_2^{2n}+2z_1^nz_2^n):(z_1^{2n}+z_2^{2n}-2z_1^nz_2^n):(2z_1^nz_2^n)\\=&(z_1^n+z_2^n)^2:(z_1^n-z_2^n)^2:2(z_1z_2)^n,\end{aligned}$$

因此上面的表中可以加上

二面体: φ_n^2, χ_n^2, ψ_2^n, $N=2n$。

由此方程的形式即可看出临界点及其重数,与我们以前找到的完全一致(见“二面体方程”一节)。

现在在这 3 种情况下,实际找出 φ, χ, ψ 的形式。我只详细介绍八面体,因为在这种情况下关系是最简单的。但即使对八面体,有时也只给出提示或结果,以便限制在简述的范围之内。希望详细了解的人,可查阅我的讨论二十面体的那本书。为简单起见。想象该八面体在 z 球面上的位置使 6 个顶点处于

$$z=0,\ \infty,\ +1,\ +\mathrm{i},\ -1,\ -\mathrm{i},$$

如图 6.17 所示。于是很容易给出代表八面体旋转,即使这 6 个点互换的 24 个线性变换。我们从顶点 0 和∞保持固定的 4 个旋转开始,它们是

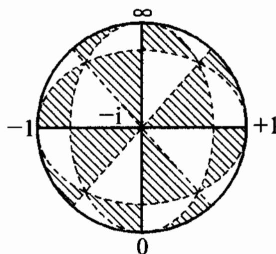


图 6.17

$$z' = \mathrm{i}^k \cdot z \quad (k = 0, 1, 2, 3)。\tag{12a}$$

然后通过代换 $z' = \frac{1}{z}$ 将 0，∞两点互换（即绕水平轴+1，−1 旋转 180°，它仍使八面体的每一个顶点转换到另一个顶点）。再作 4 个旋转(12a)，即得 4 个新变换

$$z' = \frac{\mathrm{i}^k}{z} \quad (k = 0, 1, 2, 3)。\tag{12b}$$

同样地，通过 4 个变换

$$z' = \frac{z+1}{z-1},\ \frac{z+\mathrm{i}}{z-\mathrm{i}},\ \frac{z-1}{z+1},\ \frac{z-\mathrm{i}}{z+\mathrm{i}},$$

将剩下的 4 个点 $z = 1, \mathrm{i}, -1, -\mathrm{i}$ 依次变为∞，它们显然会使八面体的 6 个顶点互换，每一次再施以变换(12a)，从而得到八面体的另外 $(4 \cdot 4 =)16$ 个变换

$$\begin{cases} z' = \mathrm{i}^k \dfrac{z+1}{z-1},\ z' = \mathrm{i}^k \dfrac{z-1}{z+1}, \\ z' = \mathrm{i}^k \dfrac{z+\mathrm{i}}{z-\mathrm{i}},\ z' = \mathrm{i}^k \dfrac{z-\mathrm{i}}{z+\mathrm{i}}, \end{cases} k = 0, 1, 2, 3。\tag{12c}$$

这样就找到了所需要的 24 个变换。通过计算，很容易证明，它们事实上将八面体的 6 个顶点互换，并形成群 G_{24}，即逐次施加任两个变换，仍然得到(12)式中的一个变换。

现在构造在八面体的顶点处为零的函数 ψ_6。点 $z=0$ 给出因式 z_1，点 $z=\infty$ 给出因式 z_2，因式 $z_1^4 - z_2^4 = 0$ 将给出单根±1, ±i，从而最后得

$$\psi_6 = z_1 \cdot z_2 (z_1^4 - z_2^4)。\tag{13a}$$

构造函数 φ_8 和 χ_{12} 要困难得多，它们分别在八面体的各面和各边的中心取零值。我们不去推导，只指出它们是[①]

① 见 *Ikosaeder*，第 54 页。

$$\begin{cases}\varphi_8 = z_1^8 + 14z_1^4z_2^4 + z_2^8, \\ \chi_{12} = z_1^{12} - 33z_1^8z_2^4 - 33z_1^4z_2^8 + z_2^{12}。\end{cases} \tag{13b}$$

不用说，在这 3 个式子中，每一个式子都有一个特定常数因子。如果 φ_8，ψ_6，χ_{12} 表示(5)式那样的规范化的形式，则在八面体方程(11)中必须插入待定常数 c_1，c_2，写成

$$w_1 : (w_1 - w_2) : w_2 = \varphi_8^3 : c_1\chi_{12}^2 : c_2\psi_6^4。$$

现在，常数 c 必须这样确定，使得两个在 z 与 w 之间的方程实际上给出的是一个方程。当且仅当

$$\varphi_8^3 - c_2\psi_6^4 = c_1\chi_{12}^2$$

对 z_1，z_2 恒成立时，这才有可能。这个关系只能由确定的常数值 c_1 和 c_2 来满足。简单的计算表明，恒等式

$$\varphi_8^3 - 108\psi_6^4 = \chi_{12}^2$$

必然成立。于是八面体方程变成

$$w_1 : (w_1 - w_2) : w_2 = \varphi_8^3 : \chi_{12}^2 : 108\psi_6^4。 \tag{14}$$

由于式子 φ，χ，ψ 按上述给出，故此方程确实将点 0，1，∞分别映射于八面体各面的中心、各边的中点和各顶点，并具有相应的重数。而且，八面体的 24 个变换(12)式将它变为自身，因为它们把式子 φ，χ，ψ 的每个零点仍变为它们自身，同时使每个式子改变一个乘数因子。计算表明，取商后，这些因子将会消失。

剩下的是证明方程(14)确实将 z 球面上有阴影或无阴影的三角形保角地映射成 w 的前半或后半球面。我们知道，w 的实子午线的点 0，1，∞对应于每个三角形的 3 个顶点，但对每个 w 值，方程有 24 个 z 的根。因为这些根必然分布在 24 个三角形内，在一个三角形内，w 最多只能取一次给定的值。只要能证明在三角形的 3 个边界上 w 保持是实的，就很容易证明存在一一对应映像，将三角形的每个边变为实

w 子午线的一段;也存在一个类似的映像将三角形的整个内部变换为相应的半球,且映像是不反转角度的定向的保角映射。利用函数$w(z)$的连续性和解析特性,你们自己也可以得出这些结论。我只指出关键的一步,即证明 w 的实线映射为三角形的边界。

有一个比较方便的证明法,就是先证明由八面体分割而产生的所有大圆对应的 w 是实的。这些大圆是:通过八面体 6 个顶点中任意 4 个顶点的 3 个互相垂直的大圆(主圆;图 6.17 的实线所示),其次是对应于各面高线并将主圆间夹角平分的 6 个圆(辅圆;图 6.17 的虚线所示)。通过相应于八面体的代换,可以将每一个主圆变换成另一个主圆,也可以将每个辅圆变换成另一个辅圆。因此只要能证明,在一个主圆和一个辅圆上的每一点,函数 w 取实数值就足够了,因为它在其他的圆上必然取同样的值。z 球面的实数子午圆是主圆之一,据(14)式,在此圆上的值为

$$w=\frac{w_1}{w_2}=\frac{\varphi_8^3}{108\psi_6^4}。$$

因 φ 和 ψ 是 z_1 和 z_2 的实多项式,w 当然也是实的。从辅圆中,选一个通过 0 和∞且与实子午圆成 45°角的圆,其上 $z=e^{\frac{i\pi}{4}}\cdot r$, r 从$-\infty$到$+\infty$取遍实数。在此圆上,$z^4=e^{i\pi}\cdot r^4=-r^4$ 是实的。因在(13)式内的 φ_8 和 ψ_6 的四次幂中只含 z_1 和 z_2 的四次幂,故上式表明 w 是实的。

由此即得证明:方程(14)事实上将 w 半球或其上的黎曼面保角映射到 z 球面的对应于八面体的三角形划分上,这样和以前的例子一样完整地从几何上求得由此方程所建立的 z 和 w 之间的依赖关系。

四面体和二十面体可按同样方法处理,我只叙述其结果。和前面一样,这些结果是在 z 球面的划分具有最简单的位置下得到的。四面体方程①是:

---

① 参阅 F·克莱因著《四面体》,第 51 页及第 60 页。

$$w_1 : (w_1 - w_2) : w_2 = (z_1^4 - 2\sqrt{-3}z_1^2z_2^2 + z_2^4)^3$$
$$: -12\sqrt{-3}[z_1z_2(z_1^4 - z_2^4)]^2$$
$$: (z_1^4 + 2\sqrt{-3}z_1^2z_2^2 + z_2^4)^3;$$

二十面体的方程①是：

$$w_1 : (w_1 - w_2) : w_2 = [-(z_1^{20} + z_2^{20}) + 228(z_1^{15}z_2^5 - z_1^5z_2^{15})$$
$$- 494z_1^{10}z_2^{10}]^3$$
$$: -[(z_1^{30} + z_2^{30}) + 522(z_1^{25}z_2^5 - z_1^5z_2^{25})$$
$$- 10\,005(z_1^{20}z_2^{10} + z_1^{10}z_2^{20})]^2$$
$$: 1\,728[z_1z_2(z_1^{10} + 11z_1^5z_2^5 - z_2^{10})]^5。$$

即：这些方程相应地将 w 半球保角地映射到相应于四面体和二十面体的 z 球面的有阴影和无阴影的三角形上。

5. 关于正规方程的解

现在考虑那些我们已讨论过的方程的某些公共性质，我们将称这些方程为正规方程。

首先要注意，我们的所有正规方程，其性质之所以极其简单，是由于把它们变为其自身的线性变换的数目正是其次数，即其所有的根都是其某一个根的线性函数；且其球面划分已将一切需要考虑的关系用十分明显的几何图像表现出来。如果结合二十面体方程来提出某一个问题，那么许多事情会显得十分简单，否则如此高次方程通常会把问题搞得很复杂。

设实值 w 被给定在 w 的实子午线的 $(1, \infty)$ 段内（图 6.18），我们要问一下当 $w = w_0$ 时二十面体方程的 60 个根 z 的问题。我们的映射理论立即告诉我们：z 球面的 60 个三角形的某一边界（图 6.16 的实线）上都必各有一个根。这样，就提出了方程式论中的所谓根的

---

① 参阅 F·克莱因著《四面体》，第 56 页及第 60 页。

分离问题。进行繁重的根的数值计算，通常必须先找出每个根所属的区间。但我们也能立即说出有多少根是实的。如果考虑到上面所给的二十面体方程的形式隐含着二十面体在 z 球面上的位置①，使得实子午线包含有(a)、(b)、(c)3 类顶点中某一类的 4 个顶点，则可推出(图 6.15 和图 6.16)有 4 个用实线画的三角形的边位于实子午线上，从而得出正好有 4 个实根。如果 w 在实子午线的另外两段上，则这个结论也是对的。这样的话，对 0，1，∞以外的每个实数 w，二十面体方程有 4 个实根、56 个虚根；对 $w=0, 1, \infty$，也有 4 个不同实根，但都是重根。

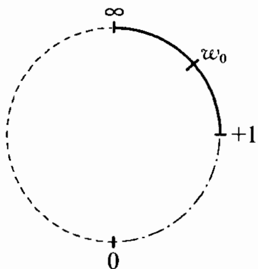


图 6.18

现在讲一讲正规方程根的实际的数值解法。由于只需要计算一个根，其他根可由线性变换求得，因此可以得到很大的好处。但要提醒一下，根的数值计算实际上是分析问题，而不是代数问题。因为，在一般情况下的根是无理数，在逼近它时必须用到无限的过程。

我只详细讲一个最简单的例子，即纯方程

$$w = z^n。$$

这里又直接涉及中学的数学内容，因为中学里也讲这个方程，即 $\sqrt[n]{w}$ 的计算，至少对小的 n 和实数值 $w=r$，中学里是要讲的。计算平方根和立方根的方法，你们在学校里已学过，本质上依赖于下面的步骤。首先决定被开方数 $w=r$ 在自然数 1，2，3，…的平方或立方数序列中所占有的位数，然后利用十进制记数法，用有关区间的十分之一作同样试验，然后用百分之一，等等。用此方法当然可以得到所需的准确度。

不过，我愿使用一个更合理的作法，此作用不仅适用于任意整数

① 参阅克莱因著《四面体》一书，第 55 页。

值 n，也适用于任意复数值 w。因为只需要确定方程的一个解，就可以专找与实数轴夹角在$\frac{2\pi}{n}$以内的那个值$z=\sqrt[n]{w}$。将前述基本方法推广，先将该角分成 ν 等分（图 6.19 中 $\nu=5$），以原点为中心，半径为 $r=1, 2, 3$，作一族同心圆，与等分角的射线相交。在选定 ν 后，即可用此法找到所有标在所述角区域内的点：

$$z=r\cdot e^{\frac{2i\pi}{n}\cdot\frac{k}{\nu}}\quad\begin{pmatrix}k=0, 1, 2, \cdots, \nu-1\\r=1, 2, 3, \cdots\end{pmatrix},$$

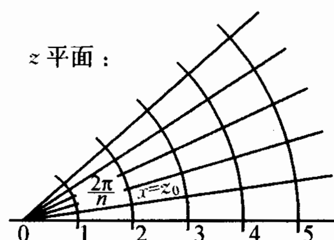


图 6.19

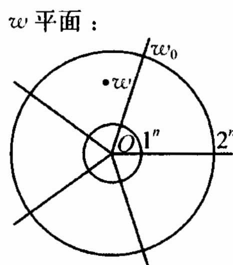


图 6.20

且能在 w 平面内立即标出对应的 w 值

$$w=z^n=r^n e^{2i\pi\cdot\frac{k}{\nu}}。$$

这些就是由半径为 1^n，2^n，3^n，… 的圆和实数轴成角 0，$\frac{2\pi}{\nu}$，$\frac{4\pi}{\nu}$，…，$\frac{(\nu-1)2\pi}{\nu}$ 的射线所组成的、覆盖整个 w 平面的网络的格点（图 6.20）。设 w 位于网络的一个眼的内部或周界上，并设 w_0 为最靠近它的格点。我们知道$\sqrt[n]{w_0}$ 的一个值 z_0 是 z 平面上的一个格点，因此所求的值应为

$$z=\sqrt[n]{w}=\sqrt[n]{w_0+(w-w_0)}=\sqrt[n]{w_0}\sqrt[n]{1+\frac{w-w_0}{w_0}}$$

$$=z_0\left(1+\frac{w-w_0}{w}\right)^{\frac{1}{n}}。$$

将右端按二项式定理展开（我们假定这定理是已知的，因为我们现在事实上已经进入分析的领域）：

$$z = z_0\left\{1+\frac{1}{n}\left(\frac{w-w_0}{w_0}\right)+\frac{1-n}{2n^2}\left(\frac{w-w_0}{w_0}\right)^2+\cdots\right\}。$$

如果把它看做是解析函数$\sqrt[n]{w}$的泰勒展开式，并把它应用于以 w_0 为圆心的、并通过最近奇异点的圆，立即可知此级数是收敛的。因为$\sqrt[n]{w}$只有 0 和∞是奇点，因此，当且仅当 w 位于以 w_0 为圆心并通过原点的圆内，我们的展开式才是收敛的，而我们总是能在 z 平面内作出类似的网络。或许有必要把网眼作小些，即可达到此目的。但为了使收敛性更好些，即使此级数更便于数值计算，$\dfrac{w-w_0}{w_0}$ 必须足够小。这一点，通过进一步缩小格子总是能办到的。这对根的数值计算是一个很有用的方法。

现在值得说明一下，求其他正多面体的正规方程的数值解，并无更大的困难，但我将省略证明。如果从映射两个相邻三角形到 w 球面出发，把同样的方法应用到我们的正规方程上，则将会出现另外的级数而不是二项级数。这种级数叫超几何级数，学过数学分析的人都知道，而且也很实用。1877 年我具体地找到了这个级数①。

6. 用超越函数将正规无理性单值化

现在要介绍另一种解正规方程的方法，其特征是系统地使用超越函数。用这种方法就不必在一个已知解的邻近作级数展开，相反可以一下子给出全部数对(w, z)，使 w 和 z 都成为一个辅助变量的单值解析函数，以满足方程，或用我们的话来说，即将方程所确定的无理性单值化。如果所用函数只是容易列表的函数，或已有数值表

① 见 *Weitere Untersuchungen Oüber das Ikosaeder*，《数学年刊》，第 12 卷，第 515 页。也可参阅 F · 克莱因 *Gesammelte Mathematische Abhandlungen*，第 2 卷，第 331 页及以后各页。

可查,那么不需要进一步计算就可得到方程的数值解。我之所以更愿意与超越函数联系起来讨论这个问题,是因为它在中学教学中有时起到一定的作用,可是在中学里,这个问题仍然常有模糊不清甚至神秘的色彩。原因在于:尽管现代复变函数理论已提供了非常明确的方法,但人们仍然墨守传统的不完全的概念。

我将首先把这些总的提示应用于纯方程。甚至在中学里,也已经用对数来求实正数 r 的正根 $z^n = r$。我们把方程写成 $z = e^{\frac{\log r}{n}}$,其中 $\log r$ 取正主值。首先用对数表求 $\log r$,然后反过来查出对应于 $\frac{\log r}{n}$ 的数即为 z。通常用 10 而不用 e 为底。此解立即可扩充到复数。令 x 等于一般的复对数 $\log w$,使之满足方程

$$z^n = w。$$

于是可得到作为 x 的单值解析函数 w 和 z

$$w = e^x, \quad z = e^{\frac{x}{n}}。$$

根据稍后将仔细研究的函数 $x = \log w$ 的多值性,对同样的 w 可得到 n 个精确值 z。我们称 x 为均匀化变量。

因为对数表中只包含实数的实对数,显然不能立即读出给定解的值。但是,应用对数的简单性质,即可以把它化为用三角函数表来计算,这是人人都会做的。如果令

$$w = u + iv = \left|\sqrt{u^2+v^2}\right| \left(\frac{u}{|\sqrt{u^2+v^2}|} + i\frac{v}{|\sqrt{u^2+v^2}|}\right),$$

第一个因子是正实数,因此有实对数,而第二个因子绝对值为 1,因而有纯虚对数 $i\varphi$(即第二个因子等于 $e^{i\varphi}$)。由此从公式

$$\text{(a)}\ \frac{u}{\sqrt{u^2+v^2}} = \cos\varphi,\ \frac{v}{\sqrt{u^2+v^2}} = \sin\varphi$$

即可求得 φ。于是 $x = \log w = \log\left|\sqrt{u^2+v^2}\right| + i\varphi$,方程的根则为

$$z = e^{\frac{x}{n}} = e^{\frac{1}{n}\log\left|\sqrt{u^2+v^2}\right|} \cdot e^{\frac{1}{n}i\varphi},$$

即有

$$(b)\ z = \sqrt[n]{u+iv} = e^{\frac{1}{n}\log\left|\sqrt{u^2+v^2}\right|}\left(\cos\frac{\varphi}{n} + i\sin\frac{\varphi}{n}\right)。$$

因为 φ 之值有一个 2π 的倍数不能唯一确定，所以这个公式给出了所有 n 个根。借助于普通的对数和三角函数表，首先可由(a)求得 φ，然后可由(b)求得 z。从复数的对数求得“三角解”，这个方法是完全自然的。然而，假设这些都不知道，而像中学里的做法来求这个三角解，就必然显得十分奇怪难懂了。

有时必须找非实数根，例如，在中学教学中，三次方程的所谓卡丹解法就必须求这种根。关于卡丹解法，我想补入一些注解。如果方程以退化形式

$$x^3 + px - q = 0 \tag{15}$$

给出，则卡丹公式指出，方程的 3 个根 x_1，x_2，x_3 可由如下表达式给出

$$x = \sqrt[3]{\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}。\tag{16}$$

因为每个立方根有 3 个值，所以这个表达式总共将给出一般不相同的 9 个值。要定出 x_1，x_2，x_3，需要一个条件，即任意两个所用的立方根之积为 $\frac{-p}{3}$。如果用众所周知的方法将系数 p，q 代之以 x_1，x_2，x_3 的对称函数，并注意到由于 x^2 的系数为 0，即 $x_1 + x_2 + x_3 = 0$，则得

$$\frac{q^2}{4} + \frac{p^3}{27} = -\frac{(x_1 - x_2)^2(x_2 - x_3)^2(x_3 - x_1)^2}{108}。$$

由此可见，解的公式中平方根内的数是方程的判别式，但相差一个负

号。当所有根是实数时,它是负的,但当一根是实的,另两个是共轭复根时,它就是正的。这正是三次方程中看来最简单的情况,即当3个根是实的时,用卡丹公式必须求负数的平方根,从而对虚数求立方根。

在高斯用平面来解释复数之前的250年,当人们还远远不懂复数理论的时候,求实解的过程必须通过复数,这在中世纪的代数学家看来想必是不可思议的。当时人们谈到过三次方程不可约情况("Casus irreducibilis"),并说卡丹公式在这时未能给出一个合理的可用的解。可是后来发现恰恰是在这种情况下可使三次方程与角的三等分之间建立起一种简单的关系,并且由此得到一个实的"三角解",以代替有缺陷的卡丹公式。这时人们才相信已有了某种新的发现,这种发现和旧公式没有联系。不幸的是,即使在今天的初等数学教学中,偶尔还有人采取这种立场。

与此观点相反,我想强调指出,这种三角解正是应用刚才讨论过的方法来求复数根。因此,当被开立方的数是复数的时候,对卡丹公式加以变换,以便用中学里学实数开方时的同样方便的方法进行数值计算,就能很自然地求出三角解。所以设

$$\frac{q^2}{4}+\frac{p^3}{27}<0,$$

如果 q 是实的,则 p 必然是负数。于是将(16)式中的第一个立方根写成

$$\sqrt[3]{\frac{q}{2}+\mathrm{i}\left|\sqrt{-\frac{q^2}{4}-\frac{p^3}{27}}\right|},$$

即知它的绝对值(即被开方数的值 $\sqrt{\frac{-p^3}{27}}$ 的正立方根)等于 $\left|\sqrt{\frac{-p}{3}}\right|$;但因它与第二个立方根之积等于 $\frac{-p}{3}$,所以第二个立方根必是它的共轭复数,两者之和即三次方程的解,正是其实数部分的两

倍,即

$$x_1, x_2, x_3 = 2R\left(\sqrt[3]{\frac{q}{2}+\mathrm{i}\left|\sqrt{-\frac{q^2}{2}-\frac{p^3}{27}}\right|}\right)。$$

现在采用前面对复数进行计算的一般方法。在提出绝对值之后,三次根式内的值可写成

$$\left|\sqrt{-\frac{p^3}{27}}\right|\left\{\frac{\frac{q}{2}}{\left|\sqrt{-\frac{p^3}{27}}\right|}+\mathrm{i}\frac{\left|\sqrt{-\frac{q^2}{4}-\frac{p^3}{27}}\right|}{\left|\sqrt{-\frac{p^3}{27}}\right|}\right\},$$

并由方程式

$$\cos\varphi=\frac{\frac{q}{2}}{\left|\sqrt{-\frac{p^3}{27}}\right|}\quad \sin\varphi=\frac{\sqrt{-\frac{q^2}{4}-\frac{p^3}{27}}}{\left|\sqrt{-\frac{p^3}{27}}\right|}$$

确定角 φ。因为 $\left|\sqrt{-\frac{p^3}{27}}\right|$ 的正立方根是 $\left|\sqrt{-\frac{p}{3}}\right|$,因而我们的立方根取形式

$$\left|\sqrt{-\frac{p}{3}}\right|\cdot\left(\cos\frac{\varphi}{3}+\mathrm{i}\sin\frac{\varphi}{3}\right)。$$

注意到 φ 可以确定到相差 2π 的倍数,故有

$$x_k=2\left|\sqrt{-\frac{p}{3}}\right|\cos\frac{\varphi+2k\pi}{3}\quad(k=0,1,2),$$

而这是三角解的常用形式。

我想借此机会就所谓"不可约情况"这个用语做一点说明。这里所用的"不可约",与现今通用的、本讲义也常用的同一用语意义完全

不同。这里的含义是三次方程的解不能由实数的三次方根推出。这根本不是现在这个词的含义。你看,用语不恰当,再加上对复数的普遍惧怕,至少在这个领域内造成了大量的误解。我希望我的话至少对你们能起预防作用。

现在简短地探讨一下在其他正则无理性中借用超越方程单值化的可能性。在二面体方程

$$z^n + \frac{1}{z^n} = 2w$$

中,简单地令

$$w = \cos\varphi,$$

棣莫弗公式表明,

$$z = \cos\frac{\varphi}{n} + \mathrm{i}\sin\frac{\varphi}{n}$$

满足该方程。因为所有的 $2k\pi+\varphi$ 和 $2k\pi-\varphi$ 给出同一的 w 值,对每一个 w,上面公式事实上给出了 $2n$ 个 z 值,它们可写成

$$z = \cos\frac{\varphi+2k\pi}{n} \pm \mathrm{i}\sin\frac{\varphi+2k\pi}{n} \quad (k = 0,1,2,\cdots,n-1)。$$

对于四面体、八面体和二十面体方程,这些"初等"超越函数是不够的。然而,借助于椭圆模函数,仍能求得相应的解。尽管人们认为这样的解不属于初等数学,但我想至少给出关于二十面体的公式①。这些公式与借助于椭圆函数解一般五次方程密切相关,这件事教科书里总会提及,稍后我也会作某些解释。二十面体方程式为

$$w = \frac{\varphi_{20}(z)^3}{\psi_{12}(z)^5}。$$

① 参阅《数学年刊》,第 14 卷(1878 年/1879 年),第 111 页及随后数页,或 F·克莱因的《数学论文集》,第 3 卷,第 13 页及随后数页,以及《四面体》第 131 页。

现在令 w 等同于椭圆函数理论中的绝对不变量 J，并设想 J 是周期商 $w=\frac{\omega_1}{\omega_2}$ $\left(按雅可比的符号为\frac{ik'}{k}\right)$的一个函数，即令

$$w=J(\omega)=\frac{g_2^3(\omega_1, \omega_2)}{\Delta(\omega_1, \omega_2)},$$

式中 g_2 和 Δ 分别是 ω_1 和 ω_2 的 -4 次和 -12 次的某个超越式子，这些式子具有重要作用。如果引入常用的雅可比简写式

$$q=e^{i\pi\omega}=e^{-\pi\frac{k'}{k}}$$

二十面体方程的根 z 将由下面 θ 函数的商给出

$$z=-q^{\frac{3}{5}}\frac{\theta_1(2\pi\omega, q^5)}{\theta_1(\pi\omega, q^5)}。$$

如果考虑到按第一个方程，ω 是 w 的函数，且有无穷多值，则本公式对给定的 w，事实上给出了二十面体方程的所有 60 个根。

7. 借助根式求解

正规方程理论中有一个问题是我们尚未触及的，即我们的正规方程从代数上说有没有提供实质上是新的东西，是否能分解为其他正规方程，特别是能否化为一系列纯方程？换句话说，是否能通过有限根式符号的叠置而用 w 求出这些方程的解 z？

就二面体、四面体和八面体方程而论，很容易借助代数理论证明，它们事实上能化简为纯方程。只消对二面体方程

$$z^n+\frac{1}{z^n}=2w$$

作出详细说明就行了。如果令

$$z^n=\zeta,$$

方程将变成

$$\zeta^2 - 2w\zeta + 1 = 0,$$

从而得出

$$\zeta = w \pm \sqrt{w^2 - 1}。$$

于是

$$z = \sqrt[n]{w \pm \sqrt{w^2 - 1}}。$$

这就是借助根式而求得的解。

然而另一方面，二十面体方程却没有这样借助根式的解。因此，这个方程确定了一个本质上是新的代数函数。关于这一点，我打算提供一个我最近发表的图解证明(《数学年刊》第 61 卷，1905 年)，这个证明是从熟知的二十面体函数 $z(w)$ 的函数论构造的考虑得出的。为此需要用阿贝尔的下述定理：如果一个代数方程的解能用一系列根式表达，则系列中的每个根式可以用给定方程的 n 个根的有理函数来表达。这个证明可以在每一本代数专著中找到。

现在把这个定理用于二十面体方程。如果假设它的根 z 能用系数的有理函数，即 w 的有理函数的一系列根式来表达，则系列中的每一个根式是 60 个根的有理函数

$$R(z_1,\ z_2,\ \cdots,\ z_{60})。$$

首先，能用只含 z 的有理函数 $R(z)$ 来代替这个表达式，因为所有的根能用线性变换从其中的一个根导出。现在借助于 60 个值的二十面体函数 $z(w)$ 将 $R(z)$ 转换成 w 的函数，并考虑其结果。因为在 w 平面上的每一个将 z 转回其初始值的环路必然将 $R(z)$ 也转到它的初始值，所以 $+R[z(w)]$ 只可能在 $w = 0, 1, \infty$ 有支点($z(w)$ 在此有支点)，而在每个这样的位置上循环连接的 $R[z(w)]$ 黎曼面叶数，也必然对应于 $z(w)$ 的叶数的一个约数。我们知道对应于这 3 个位置的叶数分别为 3，2，5。因此，二十面体方程根的每一个有理函数 $R(z)$，以及每一个出现在所设解中的根式，作为 w 的函数来考

虑,如果有支点的话,只会在 $w=0$, $w=1$, $w=\infty$ 处。如果有支点,那么必然有 3 叶在 $w=0$ 处,有两叶在 $w=1$ 处,5 叶在 $w=\infty$ 处连接,因为 3, 2, 5 没有 1 以外的除数。

现在来说这个结果所导致的矛盾。为此我们来检查对 $z(w)$ 所假设的表述式中最里层的根式。它的被开方式必然是一个有理数 $P(w)$。可以假设根指数是质数 p,因为在相反的情况下可以用根指数为质数的方根叠合而成,此外 $P(w)$ 不可能是 w 的某个函数 $\theta(w)$ 的 p 次幂,因为如果是这样,这个根式就是多余的了,可以直接考虑下一层真正的根式。

现在我们来看函数 $\sqrt[p]{P(w)}$ 含有什么类型的支点。比较方便的办法是将它写成齐次形式

$$P(w)=\frac{g(w_1, w_2)}{h(w_1, w_2)},$$

其中 g 和 h 是变量 w_1, $w_2\left(w=\frac{w_1}{w_2}\right)$ 的同样次数的形式。按代数基本定理,可将 g 和 h 分解为线性因子并写成

$$P(w)=\frac{l^{\alpha}\cdot m^{\beta}\cdot n^{\gamma}\cdot\cdots}{l'^{\alpha'}\cdot m'^{\beta'}\cdot n'^{\gamma'}\cdot\cdots}.$$

因为分子分母是齐次的,所以有

$$\alpha+\beta+\gamma+\cdots=\alpha'+\beta'+\gamma'+\cdots.$$

所有指数 α, β, $\cdots$, α', β', $\cdots$不能都被 p 整除,否则 P 就完全成为一个 p 次幂了。另一方面,$\alpha+\beta+\cdots-\alpha'-\beta'-\cdots$ 是等于零的,故能被 p 整除。因此,这些数中至少有两个不能被 p 整除。于是推出相应的两个线性因子的零点必然是 $\sqrt[p]{P(w)}$ 的支点,且有 p 叶在每个支点上循环地相连接。但这与前面所述的定理相矛盾,因为这个定理当然应该对 $\sqrt[p]{P(w)}$ 同等有效。上面我们数过所有可能的支点,发现其中没有两点具有同样的相遇叶数。因此假设是站不住脚的,

因而二十面体方程不能用根式求解。

本证明要依赖这样一件事，即作为二十面体方程特征的数字 3，2，5 彼此都没有公约数。如果存在公约数，例如相应于八面体的数字为 3，2，4，那么立即有可能存在在两点具有同类型支点的有理函数 $R[z(w)]$，例如在 1 和∞处均有两叶连接的有理函数。这样的话，这些函数就真正可以作为有理函数 $P(w)$的根而表示了。这样就得到了八面体、四面体(具有数 3，2，3)和二面体方程(2，2，n)用根式求解的方法。

我想借此机会指出，广大数学界所用的术语是多么跟不上数学知识的进展。“根”这个术语现今几乎到处用于两重含义：其一是指任意代数方程的解；其二是特指纯方程的解。后一含义当然可追溯到只研究纯方程的时代。到了今天，即使不说有害，至少也是不方便的。例如，说一个方程的“根”不能用根式来表述，这几乎是矛盾的。但还有一种会造成更严重误解的说法从一有代数就徘徊不去，即若一代数方程不能化为纯方程，即它不能用根式求解，就称它为“代数不可解”。这种说法和当今“代数”这个词的意义是完全矛盾的。今天，当我们说一个方程可以代数求解时，意思就是能将它化简成一串最简单的代数方程，其中解对参数的依赖性、不同根的相互关系等等，都完全可以像纯方程一样加以支配，这些方程不一定非纯方程不可。我们说二十面体方程是代数可解的，就是这个意思，因为我们的讨论表明，能用符合所有上述要求的方式建立二十面体方程的理论。这个方程不能用根式求解，倒是值得注意，因为这表明，还有其他一些过去所谓代数不可解的方程所要竭力化简成的(即要完全求解的)适当的正规方程，可能就是这个方程。

上面这段话把我们引到本章的最后一节，我们将概括介绍一下这种化简概念。

8. 化简一般方程为正规方程

我们知道，下面的化简是可能的：

一般三次方程可化为 $n=3$ 的二面体方程。

一般的四次方程可化为四面体或八面体方程。

一般的五次方程可以化为二十面体方程。这个结果是有数学史以来始终起着重大作用,并在大多数相去甚远的现代数学领域中具有决定性影响的正多面体理论的一个最新的辉煌成就。

为了表明这个论断的意义,将对三次方程作较详细的介绍,但对公式不给出完整的证明。再拿化简形式的三次方程作例子

$$x^3+px-q=0。 \tag{17}$$

用 x_1, x_2, x_3 表示其解,设法找到它们的一个有理函数 z,使得用所有可能的 6 个方法互换 x_i 时,该函数经受 $n=3$ 的二面体方程的 6 个线性变换。z 应取的值为

$$z,\ \varepsilon z,\ \varepsilon^2 z,\ \frac{1}{z},\ \frac{\varepsilon}{z},\ \frac{\varepsilon^2}{z}(\text{其中 } \varepsilon=e^{\frac{2i\pi}{3}})。$$

很容易看到

$$z=\frac{x_1+\varepsilon x_2+\varepsilon^2 x_3}{x_1+\varepsilon^2 x_2+\varepsilon x_3} \tag{18}$$

满足这些条件。通过所有 x_k 的互换,这个量的二面体函数 $\frac{z^3+1}{z^3}$ 必然保持不变,因为 z 的这 6 个线性变换使它保持不变。因此,根据著名的代数定理,它必然是方程(17)的系数的有理函数。计算表明

$$z^3+\frac{1}{z^3}=-27\frac{q^2}{p^3}-2。 \tag{19}$$

反之,如果解此二面体方程,且如果 z 是它的一个根,则借助于(18)式和著名的关系式

$$x_1+x_2+x_3=0,\ x_1x_2+x_2x_3+x_3x_1=p,\ x_1x_2x_3=q,$$

通过 z, p 和 q 的有理式表达出 x_1, x_2, x_3 的值。据此找到

$$\begin{cases} x_1 = -\dfrac{3q}{p} \cdot \dfrac{z(1+z)}{1+z^3}, \\ x_2 = -\dfrac{3q}{p} \cdot \dfrac{\varepsilon z(1+\varepsilon z)}{1+z^3}, \\ x_3 = -\dfrac{3q}{p} \cdot \dfrac{\varepsilon^2 z(1+\varepsilon^2 z)}{1+z^3}。 \end{cases} \tag{20}$$

因此,只要解出二面体方程(19),(20)式立即给出三次方程(17)的解。

用同样的方法可以化简一般的四次和五次方程式。当然,这些方程会长些,但原则上并不更困难。唯一新的东西是上面由方程的系数有理地表达的正规方程的参数 w $\left(2w=-27\dfrac{q^2}{p^2}-2\right)$,现在可能包含平方根。你们会在我的关于二十面体的讲义的第二部分找到关于五次方程的全部理论。书中不仅对公式进行了计算,而且对出现这些方程的实质的理由都作了解释。

最后,让我对这些研究与三次、四次与五次方程理论通常的讲法的关系说几句话。首先,如果利用借助根式表达的二面体、四面体和八面体方程的解,那么通过适当的化简,就能用我们的公式求得三次和四次方程的一般解。对于五次方程,大多数教材都不幸地只限于介绍方程不能用根式求解的否定结果,再加上含混地暗示可以通过椭圆函数求解,而严格地说应是用椭圆模函数求解。我采取了不同的办法,因为上述办法没有促进对情况的真正了解,而只作了单方面的对比和暗示。从前面所作的概述,先用代数的语言然后用分析的语言,我们可以说:

(1) 一般的五次方程确实不可能简化成纯方程,但可以简化为作为最简单的正规方程的二十面体方程。这是它的代数解的实质问题。

(2) 另一方面,二十面体方程可以通过椭圆模函数求解。就数值计算的目的来说,这完全类似于用对数来解纯方程。

这就提供了五次方程问题的完整解答。应记住的是，当通常的路径不能获得成功时，不应满足于确定不可能性，而应激励自己去寻找新的和比较走得通的道路。这样说来，数学思想是无止境的。如果有人对你说数学的推论已不能超越某一点，你可以确信，真正有兴趣的问题正是由此开始。

最后也许可以说，这些理论并非以五次方程为限。相反，只要利用正多面体的更高维的类比，就能对六次和更高次方程作出类似的研究。如果你们感兴趣的话，可以读我的文章：《关于五次及六次一般方程解》[1]。P·戈登(P. Gorden)及A·B·科布尔(A. B. Coble)结合我那篇文章也对此问题进行了成功的研究[2]，在后者的文章中还进行了某种程度的简化[3]。

① 刊于《数学月刊》，第129卷(1905年)，第151页，并参阅《数学年刊》，第61卷(1905年)，第50页。

② 戈登，见《数学年刊》，第61卷(1905年)，第50页，以及第68卷(1910年)，第1页。科布尔，见《数学年刊》第70卷(1911年)，第337页。

③ 并可参阅克莱因著《数学论文集》，第2卷，第502—503页。

第三部分

分　　析

在这个讲习班的后半学期，将选择分析中我们认为是重要的某些章节，像算术与代数一样加以讨论。我们所要讨论的最重要的部分是初等超越函数，即对数函数和指数函数，还有三角函数，因为它们在中学教学中起着重要的作用。现在就从对数函数和指数函数讲起。

第七章
对数函数与指数函数

我先简短地回顾一下大家所熟悉的中学课程，直到它后面所谓代数分析开始之处为止。

§7.1 代数分析的系统讨论

人们是从形式为 $a=b^c$ 的指数幂开始的，其中，指数 c 首先是正整数，逐步扩展到 c 是负整数，然后到分数，最后如果情况允许，就推广到无理数。在这个过程中，根的概念作为一个特殊的幂而出现。我不去细说这个发展而只提出乘法规则

$$b^c \cdot b^{c'} = b^{c+c'}。$$

它将两数的乘法化为指数的加法。这种化简法，你们知道是计算的基础。其所以可能，是在于乘法与加法的基本规律十分类似。这两种运算都是可交换的以及可结合的。

幂的逆运算产生了对数。量 c 称为以 b 为底的 a 的对数

$$c=\log_b a。$$

到了这一步，许多本质困难就出现了，这些困难通常都不加解释地放了过去。因此，我打算把这一步特别讲得清楚一些。为方便起见，用 x 和 y 代替 a 和 c，因为我们希望研究这两个数之间的相互依赖关系。于是，基本方程变为

$$x=b^y, \quad y=\log_b x。$$

我们首先注意，b 总是假设为正数。如果 b 是负的，则对整数值 y，x 会交替取正负值，而当 y 是分数时，则 x 甚至会是虚数，使得数对 (x, y) 的总体不可能形成连续曲线。但即使 $b>0$，如果不作一些规定，仍无法进行，而这些规定看起来颇有些随心所欲。因为，如 y 是有理数，例如 $y=\frac{m}{n}$，m，n 是互质的正整数，则 $x=b^{\frac{m}{n}}$ 定义为 $\sqrt[n]{b^m}$，当 n 为偶数时，即使限于实数，我们也得处理两个值。习惯上规定 x 只取正根，即所谓主根。

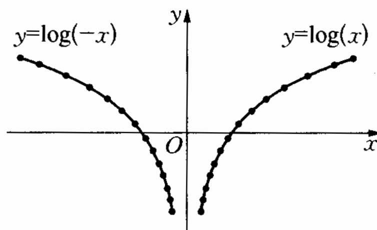


图 7.1

如果允许我早一步使用熟悉的对数 $y=\log x$ 的图像(图 7.1)，那么就可以看出，上述规定的适宜性绝非不言而喻。如果 y 取遍有理数的稠密集，则对应于横坐标为正主值 $x=b^y$ 的点，将在曲线上组成一个稠密集。现在，如果 y 的分母是偶数，我们还应该标出对应于负 x 值的点，就又得到一个点集。有人也许会说这个点集只有一半稠密。但是在对称于 y 轴的曲线 $y=\log(-x)$ 上仍然是稠密的。如果现在允许 y 取所有实数，包括无理数，那么当然就不能马上搞清楚，标在右侧的主值为什么就构成连续曲线，而标在左侧的负值是否也同样构成连续曲线。稍后将看到，这一点只有用函数论的更深奥的知识才能说清楚，而这并不在中学生的掌握范围以内。就是因为这个原因，不打算在中学里作完整的解释。人们宁愿采纳一个权威的规定，而这对中学生又是颇有说服力的，即必须取 $b>0$ 且必须取 x 的正主值，其他情况都不许提。然后就给出一个定理：对数是只对正的自变量才有定义的单值函数。

理论知识讲到了这里，就把对数表交到学生手中，要求他们学会在实际计算中使用。可能还有一些学校——我上学时一向如此——很少谈及，甚至根本不谈及这些表是怎么造的。这种不问究竟的态

度是可鄙的实用主义，它藐视任何一种比较高级的教学原则，对它必须严厉谴责。今天，对数的计算或许在大多数情况下都讨论了，而且在许多学校，为此还教了自然对数的理论和它的级数展开。

关于前一点，你们都知道，自然对数系统的底是数

$$e = \lim_{n \to \infty}\left(1 + \frac{1}{n}\right)^n = 2.718\,281\,8\cdots,$$

e 的这个定义，通常都放在大部头的分析教科书的最开始处，这是模仿法国的教法，而丝毫不讲它的来由，这样就丢掉了真正有价值的、能促进理解的部分，即不解释为什么恰好用这样特别的极限做底，为什么由此导出的对数称为自然对数。也同样突如其来地引入级数展开式，形式地假设，存在一个展开式：

$$\log(1 + x) = a_0 + a_1 x + a_2 x^2 + \cdots。$$

借助于对数的已知性质计算系数 a_0，a_1，…，或许还对 $|x| < 1$ 证明其收敛。但又不解释，甚至不怀疑像对数这种按中学的定义是如此凑合起来的函数，为什么会有级数展开式。

§7.2　理论的历史发展

如果我们希望找到学校里不讲的那些基本联系，并弄清为什么这种表面上任意的约定却能导致合理结果的深一层原因，简而言之，如果希望进一步全面了解对数的理论，最好是大体上遵循它的历史发展。你们会看到历史情况和上面提到的讲法截然不同。上面所述的讲法，可以说，是从最不利的地位对历史发展所作的投影。

我们首先讲一下 16 世纪的一位德国数学家斯瓦比亚人迈克尔·史蒂菲尔(Michael Stifel)，1544 年在纽伦堡出版了他的一本书，书名叫 *Arithmetica Integra*。那个时期是我们现在的代数学的初创时期，也就是我们曾提到的卡丹的那本书出版(也是在纽伦堡)

的前一年。这本书以及我以后还会提到的大多数书，我都可以拿给你们看。因为我们哥廷根大学的图书馆有非常完整的藏书。你们会看到，史蒂菲尔的这本书第一次用到以任意有理数为指数的幂运算，并特别强调了乘法规则。事实上，迈克尔·史蒂菲尔也可以说在一定意义上提出了第一份对数表，当然是很初步的。它只包含 2 的从 -3 到 6 的整数指数，对应的幂从 $\frac{1}{8}$ 到 64。史蒂菲尔看来懂得这一发展的意义，我们在一开头就讲到这种发展。他宣称，对这些奇异的数字关系可以写出一本书。

但为了使对数真正可用于实际计算，史蒂菲尔还缺少一个重要的手段，即十进制小数。只是在 1 600 年后它变得十分普遍时，才有可能建造实数对数表。第一份表是苏格兰人纳皮尔(1550—1617年)造出来的，收在 1614 年在爱丁堡出版的一本书中，书名为 *Mirifici logarithmorum canonis descriptio*，其所引起的兴趣，可以许多人为赞颂对数表而在前言中所作的诗为证。但是纳皮尔计算对数的方法直到他逝世后的 1619 年才出版，书名为 *Mirifici logarithmorum canonis constructio*。

还有一个瑞士人约布斯特·比尔吉(Jobst Bürgi，1552—1632年)曾独立于纳皮尔计算了一个表，不过直到 1620 年才在布拉格出版，书名为 *Arithmetishe und geometrische Progresstabuln*。我们哥廷根人想必对比尔吉产生特殊的兴趣，因为他在卡斯尔生活过很长一段时间，是我们的一个同胞。总的来说，在无穷小演算发现之前，卡斯尔，特别是那里的古天文台，对于算术、天文以及光学的发展有重要的意义，就好像汉诺威成为莱布尼兹的故乡以后变成重要的城市一样。所以，我们这个近邻城市在哥廷根大学建立以前很久就对我们科学的发展起过历史性的重要作用。

追踪纳皮尔和比尔吉的思路十分有启发性。两人都从相应于整数 y 的值 $x=b^y$ 出发，寻求一个使数 x 尽可能靠近在一起的方案。他们的目的是对每个 x 找到尽可能接近的对数 y。如前面所看到

的，今天的中学教学，是通过考虑分数值 y 来达到这个目的的。但纳皮尔和比尔吉具有天才的直觉，一下子抓住了关键而避免了呈现在他们面前的困难。他们有一个简单而美好的想法：将底 b 选到接近 1，此时，b 的逐次整数幂相互很靠近。比尔吉取

$$b = 1.0001,$$

而纳皮尔则取一个小于 1 而又靠近 1 的值

$$b = 1 - 0.0000001 = 0.999999。$$

纳皮尔采用和今天不同的方法其原因在于他想到的是对三角学计算的应用，因为在三角计算中主要是与真分数（正弦与余弦）的对数打交道，故当 $b > 1$ 时它们是负数而当 $b < 1$ 时才是正数。但这两位研究者都只用到 b 的整数幂，从而完全避免了前述使我们感到为难的多值性。

现在我们按比尔吉的系统来计算两个相邻指数 y 和 $y+1$ 的幂

$$x = (1.0001)^y, \quad x + \Delta x = (1.0001)^{y+1}。$$

相减之后，得

$$\Delta x = (1.0001)^y(1.0001 - 1) = \frac{x}{10^4},$$

或用 Δy 表示指数值之差（即 1），则有

$$\frac{\Delta y}{\Delta x} = \frac{10^4}{x}。\tag{1a}$$

由此得到一个求比尔吉对数的差分方程。比尔吉本人就直接用这个差分方程来计算对数表。确定了对应于一个 y 的 x 之后，通过加 $\frac{x}{10^4}$，他就求出了对应于 $y+1$ 的 x。同法可推出，纳皮尔的对数表满足差分方程

$$\frac{\Delta y}{\Delta x}=-\frac{10^7}{x}。\tag{1b}$$

要看出两个系统的密切关系，只要相应地把 y 写成$\frac{y}{10^4}$或$\frac{y}{10^7}$，即移动对数中的小数点即可。如果仍将这样得到的新数简单地表示为 y，则在两种情况下都会得到一串满足差分方程

$$\frac{\Delta y}{\Delta x}=\frac{1}{x}\tag{2}$$

的数，而在两种情况下，y 在每一步的步长分别为 0.000 1 和 −0.000 000 1。

如果为了方便而使用连续的指数曲线图像(按上面讨论的结果，我们得到的一定是连续曲线)，则将得到对应于纳皮尔和比尔吉数列的点的直观表示。这些点是内接于两条指数曲线

$$x=(1.000\,1)^{10\,000y}，和\ x=(0.999\,999\,9)^{10\,000\,000y}\tag{3}$$

之一的阶梯线的角点，其竖边的长在两系统中分别为 $\Delta y=0.000\,1$ 和 $\Delta y=0.000\,000\,1$ (图 7.2)。

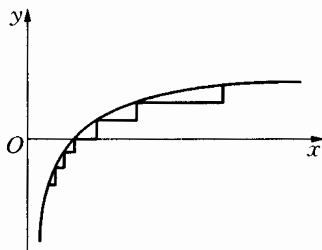


图 7.2

我们也可以得出另一个几何解释，据此可不必预先假设指数曲线而可得出获得该曲线的自然的方法。为此，只要用和的方程代替差分方程，即按方式

$$\eta=\sum\frac{\Delta\xi}{\xi}\tag{4}$$

累积起来，也就是在某种意义上积分此差分方程。在此和式中，ξ 从 1 起不连续地增加，而每一步使 $\Delta\eta=\frac{\Delta\xi}{\xi}$ 分别等于 10^{-4} 或 10^{-7}，也就是使 $\Delta\xi$ 分别按$\frac{\xi}{10^4}$与$\frac{-\xi}{10^7}$增加。在最后一步，ξ 达到值 x。人们不难

给出此过程的几何解释。为此，我们在 $\xi\eta$ 平面上画一条双曲线 $\eta=\frac{1}{\xi}$（图 7.3），并从 $\xi=1$ 开始在 ξ 轴上按每次累进 $\Delta\xi=\frac{\xi}{10^4}$（我们限于比尔吉系统）逐次描点，并在所得区间上相应竖起高为 $\frac{1}{\xi}$ 的矩形 $\left(\text{矩形面积为常数} \Delta\xi\cdot\frac{1}{\xi}=\frac{1}{10^4}\right)$。据(4)式，比尔吉对数将等于在 1 与 x 之间内接于抛物线的这些矩形面积的 10^4 项和。对纳皮尔对数也可得到类似结果。

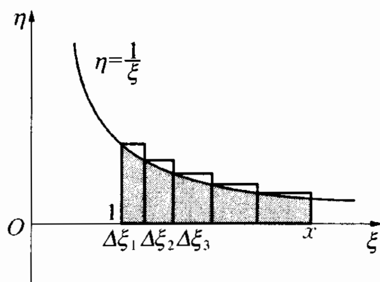


图 7.3

从这个表达式出发，如果用介于 $\xi=1$ 与 $\xi=x$ 之间的双曲线下面的面积（图中阴影部分）代替矩形面积之和，立即就得出自然对数。这表现在下述著名公式中

$$\log x=\int_1^x\frac{\mathrm{d}\xi}{\xi}。$$

历史的道路实际上就是如此。1650 年左右，当解析几何已成为数学家的普通工具，无穷小分析已在求曲边图形的面积方面取得成功时，就跨出了决定性的一步。

如果希望用自然对数的这个定义作为出发点，当然必须说明它具有将数的乘法变为对数的加法这个基本性质，或用现代的语言，必须指明由双曲线下面的面积定义的函数

$$f(x)=\int_1^x\frac{\mathrm{d}\xi}{\xi},$$

具有简单的加法定理

$$f(x_1)+f(x_2)=f(x_1\cdot x_2)。$$

事实上，如果改变 x_1 和 x_2，根据积分的定义，上式两边的增量 $\frac{dx_1}{x_1}+\frac{dx_2}{x_2}$ 和 $\frac{d(x_1 \cdot x_2)}{(x_1 \cdot x_2)}$ 是相等的。所以 $f(x_1 \cdot x_2)$ 和 $f(x_1)+f(x_2)$ 只可能相差一个常量，而取 $x_1=1$ 时即知此常量变为 0(因 $f(1)=0$)。

如果希望用这种方法确定所得到的对数的"底"，那么只需注意，将增量 $\Delta\xi=\frac{\xi}{10^4}$ 改为 $\Delta\xi=\frac{\xi}{n}$ 并让 n 变为无穷，即可将矩形面积之和转化为双曲线下的面积。这与将比尔吉数列 $x=(1.000\,1)^{10\,000y}$ 换成 $x=\left(1+\frac{1}{n}\right)^{ny}$ 并令 n 取整数值增大至无穷是一回事。按幂的一般定义，这等于说 x 是 $\left(1+\frac{1}{n}\right)^n$ 的 y 次幂。因此，说底是 $\lim\limits_{n\to\infty}\left(1+\frac{1}{n}\right)^n$ 似乎是合理的，通常在定义 e 时，正是取这个极限为出发点。有趣的是：比尔吉的底 $(1.000\,1)^{10\,000}=2.718\,146$ 和 e 的前 3 位小数是一致的。

现在我们考察一下对数理论在纳皮尔和比尔吉之后的发展。首先我要作如下说明：

(1) 我们在本书第一部分提到过的默卡多是第一个用双曲线的面积来定义对数的人。他在 1668 年的 *Logarithmotechnica* 书中，以及在发表于 1667 年及 1668 年《伦敦皇家科学院哲学院报》的一些文章中，利用我刚才以现代术语向你们解说过的同样的论据指出，$f(x)=\int_1^x \frac{d\xi}{\xi}$ 与以 10 为底的常用对数(它已在实际计算中使用)只差一个常数因子。这个常数因子就是所谓对数系统的模。而且他已经引入了"自然对数"或"双曲线对数"的名称[①]。但他的最大成就是建立了对数的幂级数，这是他通过除法展开和逐项积分从积分表示

① 见《伦敦皇家科学院哲学院报》，第 3 卷(1668 年)，第 761 页。

式中得到(至少是基本上得到)的。我在第一部分曾作为数学上的一个划时代的进步而提过此事。

(2) 在这方面,我也告诉过你们,牛顿采纳了默卡多的思想,并用两个重要的结果丰富了它,即一般二项式定理和级数的反演方法。后者出现在牛顿青年时代的一篇文章中,名为"De analysi per aequationes numero terminorum infinitas",后来曾印行,但1669年后以手稿形式散发①。在这篇文章中,牛顿第一次通过 $y = \log x$ 的默卡多级数的反演推出了指数级数

$$x = 1 + \frac{y}{1!} + \frac{y^2}{2!} + \cdots,$$

由此得到自然对数 $y = 1$ 的数

$$\mathrm{e} = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots。$$

借助于关于对数的函数方程,现在便不难证明,对每个有理数 y, x 是 e^y 的取值之一,事实上是通常的幂的定义下的正值。稍后,我们对此将作较详细介绍。于是,函数 $y = \log x$ 确实成为普通定义的以 e 为底的对数,而 e 则由级数而非由 $\lim\limits_{n\to\infty}\left(1+\frac{1}{n}\right)^n$ 给出。

(3) 泰勒在作出以他的名字命名的一般级数展开法之后,就可用方便得多的方法得出指数函数级数②。这样他就可以使用关系

$$\frac{\mathrm{d}\log x}{\mathrm{d}x} = \frac{1}{x}。$$

此式可由对数的积分定义得出,而由它又可导出逆关系

---

① 初见于1711年,引自《牛顿散篇文章集》,一卷本,洛桑,1744年。

② 这出现在他的作品 *Methodus Incrementorum*(伦敦,1715年)中,本书以后还要详说。

$$\frac{de^y}{dy} = e^y,$$

并且，作为他的一般级数的特殊情况，立即可写出指数的级数。

在前面的第一部分已经看到，继这一段多产时期之后，出现了一段批判时期，我几乎想说成是精神绝望的时期。在这段时期里，一切努力都是为了把新的结果放到一个完美的基础上，并筛选出错误的东西来。现在我们来看看欧拉和拉格朗日是用什么态度来对待指数函数与对数的！

我们从欧拉的《无穷小分析引论》[①]开始。首先要钦佩欧拉在推演中所表现出来的非凡的分析技巧，然而要同时注意到，他还没有显示出今天所要求的严格性的迹象。

欧拉把二项式定理放在他的讲法的开头：

$$(1+k)^l = 1 + \frac{l}{1}k + \frac{l(l-1)}{2!}k^2 + \frac{l(l-1)(l-2)}{3!}k^3 + \cdots,$$

其中指数 l 设为整数。然后在“引论”中已不考虑整数指数了。把这个展开式专用于

$$\left(1+\frac{1}{n}\right)^{ny},$$

其中 ny 是整数。然后他让 n 变到无穷，把这个极限过程施加于级数的每一项，想到 e 是由 $\lim\limits_{n\to\infty}\left(1+\frac{1}{n}\right)^n$ 定义的，所以得到指数级数

$$e^y = 1 + y + \frac{y^2}{2!} + \frac{y^3}{3!} + \cdots。$$

肯定地说，欧拉完全不关心每一步推导在现代意义上是否严格；特别是，级数各项的极限和是否真的等于各项和的极限。现在，如你们所

---

① 见第 7 卷第 85 页及随后部分，洛桑，1748 年，有马塞(Maser)的德文译本，1885 年柏林版，第 70 页。还可参阅《欧拉著作集》，第 8 卷，1923 年，由 F·拉迪奥等 3 人编。

知，这个指数的推导过程已是许多无穷小分析教材的一个模式，尽管随着时间的推移，各个不同步骤已讲得越来越精密，且其正确性也受到了严格性的考验。如果你们回忆起用字母 e 表示这个重要的数正是由于欧拉的功劳（在他这部书的第 90 页上写道："Ponamus autem brevitatis gratia pro numero hoc 2.718 28… constanter litteram e"），你们就能看到他的著作对这些问题的整个演变过程有多么大的影响。

我还要说，欧拉紧接着用完全类似的推导方法，导出了正弦和余弦级数。为此，他将函数 $\sin\varphi$ 展开为 $\sin\frac{\varphi}{n}$ 的幂并令 n 趋向无穷。如果将棣莫弗公式

$$\cos\varphi+\mathrm{i}\sin\varphi=\left(\cos\frac{\varphi}{n}+\mathrm{i}\sin\frac{\varphi}{n}\right)^n=\left(\cos\frac{\varphi}{n}\right)^n\left(1+\mathrm{i}\tan\frac{\varphi}{n}\right)^n$$

展开成幂级数，则 $\sin\varphi$ 的展开无非是将极限过程施加于二项式定理。

现在来考虑拉格朗日的《解析函数论》[①]。应再一次指出，最多只是偶然地考虑了收敛性问题。我在第一部分里曾说过，拉格朗日只考虑那些由幂级数给出的函数，并通过导出的幂级数形式定义微商。于是，泰勒级数

$$f(x+h)=f(x)+hf'(x)+\frac{h^2}{2!}f''(x)+\cdots$$

只不过是原来按 $x+h$ 展开的 $f(x+h)$ 的级数，再把 $x+h$ 的幂展开并在形式上重新排列而已。当然，如果希望把这个级数应用到一个给定的函数，则事先应该指明这个函数是解析的，即能展开成幂级数。

---

① 1797 年巴黎版，1881 年巴黎重印本卷 4。特别请比较第 3 章第 34 页及随后部分。

拉格朗日从研究函数 $f(x)=x^n$（其中 n 为有理数）开始，把 $f'(x)$ 确定为 $(x+h)^n$ 的展开式中 h 的系数，这个展开式的前两项他认为是已经计算出来了的。然后用同样的方法，他立刻得到 $f'(x)$，$f''(x)$，…。而 $(x+h)^n$ 的二项展开式则作为 $f(x+h)$ 的泰勒级数的一个特殊情形。此外，我要更明确地指出，拉格朗日并没有对无理指数作专门考虑，而考虑所有有理数时，又把无理指数看做是显然已经解决的。仔细想想这件事是很有意义的。因为今天最重视的就是对这类过渡类型进行严格的证明。

拉格朗日用这些结果来同样地处理函数 $f(x)=(1+b)^x$。他记下 $(1+b)^{x+h}$ 的二项式级数，把 $f'(x)$ 作为 h 的系数求出，然后按同样的方法确定 $f''(x)$，$f'''(x)$，…，最后求出 $f(x+h)=(1+b)^{x+h}$ 的泰勒级数。随后令 $h=0$，得到所需要的指数级数。

我打算结束这一段简短的历史回顾，其中当然只提到了第一流数学家的名字。现在我要指出 19 世纪发生了什么根本转折，以作为这段历史回顾的结束。

(1) 首先要指出有关无穷级数及其他无穷过程收敛的一些精确思想。领先的是高斯在 1812 年发表的“Abhandlung über die hyperometrische Reihe”（“论超几何级数”）一文（文中说 Disquisitiones generales circa seriem infinitam $1+\left(\frac{a\cdot b}{1\cdot c}\right)x+\cdots$）[①]。其后则有阿贝尔在 1826 年发表的关于二项式级数的研究报告（Untersuchungen über die Reihe $1+\frac{m}{1}x+\cdots$）[②]。与此同时，柯西在 19 世纪 20 年代初写的 *Cours d'Analyse*（《分析教程》）[③]中第一次

① *Commentationes societatis regiae Göttingiensis recentiorer*，第 11 卷(1813 年)第 1 期，第 1—46 页。德文译本《高斯文集》，第 3 卷，第 123—162 页，西蒙译，柏林，1888 年。

② 《数学月刊》第 1 卷，第 311—339 页，1826 年。*Ostwards Klassiker* 第 74 期。

③ *Première Partie*，*Analyse Algebrique*，巴黎 1821 年。有德文译本，柏林，1885 年，伊特齐格桑(Itzigsohn)译。

对级数收敛性作了一般讨论。这些早期研究的结果,对于我们所考虑的级数,其展开式有时是正确的,虽然严格的证明是十分复杂的。要详细了解现代的这类证明,我又要建议你们参考伯克哈特的《代数分析》或韦伯一韦尔斯坦因的教科书。

(2) 虽然我们以后将有机会作详细讨论,但是在此必须提到柯西对无穷小分析所起的决定性奠基作用。由比尔吉和纳皮尔在 17 世纪开创的对数理论,之所以建立在圆满的数学的严格基础之上,也是利用柯西的工作的结果。

(3) 最后必须提到复变函数理论(通常简称为函数论)的创立,这对于完全理解指数函数与对数函数是必不可少的。高斯是第一个全面研究这个理论基础的人,尽管这方面他没有发表什么文章。在 1811 年 12 月 18 日给贝塞尔的信(很迟才发表)[①]中,他十分清楚地描述与解释了在复平面上作为一个无穷多值函数的定积分 $\int_1^z \frac{dz}{n}$ 的意义。但是,独立建立复变函数理论并介绍给数学界的人,是柯西。

就我们所讨论的特定的课题而言,这些研究的结果可简述如下:通过求双曲线下面积而引出对数,其方法与其他任何数学方法一样严格,但其简单和清晰的程度则超过了其他方法。

§7.3　中学里的对数理论

值得注意的是,这个现代的发展在中学数学教学中几乎没有一点反映就被绕过去了,我经常提到这是一个罪过。尽管繁冗的代数分析有种种困难和不完善之处,但教师仍然勉强使用它,而不用很方便的无穷小演算,虽然 18 世纪对它的畏怯态度早已失去根据。其原因或许在于自 19 世纪开始以来,学校里的数学教学与数

---

① *Briefwechsel zwischen Gauss und Bessel*,奥沃斯(Auwers)编,柏林,1880 年,或《高斯文集》,第 8 卷,第 90 页,1900 年。

学进展完全脱了节。考虑到对未来的数学教师进行特殊训练正是从那个世纪早期开始的,这就更加值得注意了! 我在前言中曾呼吁对这种脱节现象加以注意,它来源已久,而且阻碍了中学教学的每一种改革。具体地说,在中学里人们很少考虑所讲的定理在大学里是否有了推广,而往往满足于今天也许够用,但不能适应以后需要的定义。一句话,欧拉的说法仍然是中学里的标准说法。反之,大学也往往丝毫不顾与中学教学的联系,只顾建立自己的体系。有时用一点简述或者加上"你们在中学里已学过"等不适当的说明对这种联系一笔带过。

有趣的是:向广大学生授课的大学教师,即教理工科大学生的教师,自动地采取了与我推荐的方法非常相似的对数介绍方法。这里让我特别提一下谢弗(Scheffer)编的《理工科大学生数学教科书》①。在该书第六、第七章中,你们可以找到有关对数函数及指数函数的极详尽的理论,与我的方案完全吻合,其后第八章则为类似的角函数理论。我鼓励你们看看这本书。它对所针对的教师是非常适用的,材料充分,而且不难读,甚至照顾到了不太聪明的学生的理解力。同时还请注意作者的高度的教学法技巧。举一个例子来说,对于一旦懂了之后需要记住的少量对数理论中的公式,作者不断提请学生注意,这样需要的时候就很容易查出。作者用这种方式鼓励学生在面对大堆新材料的情况下坚持学习下去。我还要提请你们注意,尽管谢弗肯定这个问题在中学里已经学过,但他还是在书里详细解说,因为他假定中学里学过的东西大多数已经忘了。不过谢弗没有像我现在那样想到提出对中学教学进行改革的建议。

我愿意把我在中学里如何简单而自然地介绍对数的方案再概述一遍。第一个原则是求已知曲数的积分而导出新的函数,这是适当的出发点。我已经说过,这不仅符合历史情况,也与高等数学

① 谢弗,莱比锡,1905年,1921年第5版。

中椭圆函数的处理相一致。遵循这个原则，可以从双曲线 $\eta = \frac{1}{\xi}$ 出发，将 x 的对数定义为在此曲数下介于坐标 $\xi = 1$ 和 $\xi = x$ 之间的面积（图 7.4）。如允许末端坐标改变，就很容易看到其面积如何随 ξ 变化，并近似画出曲数 $\eta = \log\xi$。

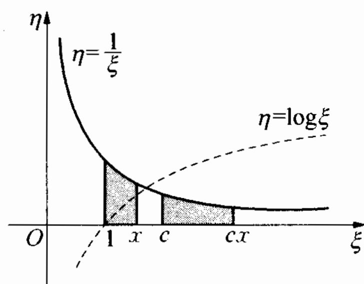


图 7.4

现在，为了简单地求得对数的函数方程，可以从关系式

$$\int_1^x \frac{d\xi}{\xi} = \int_c^{cx} \frac{d\xi}{\xi}$$

出发。此式可以对积分变量施以变换 $c\xi = \xi'$ 而得到。这意味着在坐标 1 与 x 之间的面积和坐标 c 与 cx 之间的面积一样，就离开原点来讲，后者 c 倍于前者。当我们沿曲线下 ξ 轴移动时，若按同样的比例降低高度而伸长宽度，则其面积保持不变。注意到这一点，就可以从几何上搞清楚上面所说的话。由此即可推出加法定理

$$\int_1^{x_1} \frac{d\xi}{\xi} + \int_1^{x_2} \frac{d\xi}{\xi} = \int_1^{x_1} \frac{d\xi}{\xi} + \int_{x_1}^{x_1 \cdot x_2} \frac{d\xi}{\xi} = \int_1^{x_1 \cdot x_2} \frac{d\xi}{\xi}。$$

我非常希望有人把这个方案拿到中学里去试一试。当然，如何具体实施，还得请有经验的中学教师来决定。在梅伦中学课程要求中，我们不曾冒昧地把这个方案当作标准方案提出来。

§7.4　函数论的观点

最后看看现代函数理论如何处理对数。我们会发现，早先讨论中所遇到的所有困难都完全解除了。从现在起，将用复变量 $w = u + iv$ 和 $z = x + iy$ 代替 y 和 x。

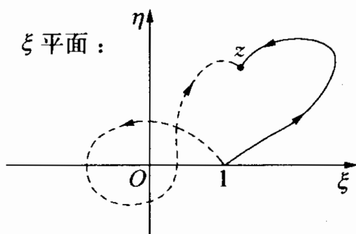


图 7.5

（1）对数定义为积分

$$w=\int_1^z \frac{\mathrm{d}\xi}{\xi}, \tag{1}$$

其中积分路径为 ξ 平面上连接 $\xi=1$ 到 $\xi=z$ 的任意曲线(图 7.5)。

（2）按积分路线围绕原点 O，1，2，…次。此积分具有无穷多个值，所以 $\log z$ 是一个无穷多值函数。如果沿负实轴切开平面并约定积分路线不越过此切口，则可定出一个确定的值，即主值$[\log z]$。当然，选择从上面还是从下面达到负实数，还具有任意性。按照这种选择，对数的虚数部分可以是 $+\pi\mathrm{i}$ 或 $-\pi\mathrm{i}$ 的主值加上 $2\pi\mathrm{i}$ 的任意倍数，即得对数的通值

$$\log z=[\log z]+2k\pi\mathrm{i} \quad (k=0,\pm1,\pm2,\cdots)。 \tag{2}$$

（3）从 $w=\log z$ 的积分定义推出其反函数 $z=f(w)$ 满足微分方程

$$\frac{\mathrm{d}f}{\mathrm{d}w}=f。 \tag{3}$$

由此可立刻写出 f 的幂级数

$$z=f(w)=1+\frac{w}{1!}+\frac{w^2}{2!}+\frac{w^3}{3!}+\cdots。$$

因为此级数对任何有限的 w 均收敛，可立即推出反函数是仅在 $w=\infty$ 有奇点的单值函数，它是一个超越整函数。

（4）和实变量一样，可从积分定义推出加法定理。由此，得反函数方程

$$f(w_1)\cdot f(w_2)=f(w_1+w_2)。 \tag{4}$$

类似地，从(2)式得出：

$$f(w+2k\pi\mathrm{i}) = f(w) \quad (k = 0, \pm 1, \pm 2, \cdots), \qquad (5)$$

即 $f(w)$是以 $2\pi\mathrm{i}$ 为周期的周期函数。

(5) 如令 $f(1) = \mathrm{e}$，则从(4)式推出，对每一个有理数值 $w = \frac{m}{n}$，$f(w)$是 $\sqrt[n]{\mathrm{e}^m}$的 n 个值之一，根式的定义和平常一样，即

$$f\left(\frac{m}{n}\right) = \sqrt[n]{\mathrm{e}^m} = \mathrm{e}^{\frac{m}{n}}。$$

我们将采用习惯的记号并用 $\mathrm{e}^w = \mathrm{e}^{\frac{m}{n}}$ 表示 $f(w)$的这一个值，故 e^w 是一个完全确定的单值函数，事实上是由方程(3)给出的那个函数。

(6) 对任意底数 b 的幂 b^w，在最一般的意义下，应了解为何种类型的函数。当然，必须采用这样的规定，使指数规律得到满足。为了建立 b^w 和刚定义的函数 e^w 之间的联系。令 b 等于 $\mathrm{e}^{\log b}$，其中 $\log b$ 取无穷多个值

$$\log b = [\log b] + 2k\pi\mathrm{i} \quad (k = 0, \pm 1, \pm 2, \cdots)。$$

于是得到

$$b^w = (\mathrm{e}^{\log b})^w = \mathrm{e}^{w\log b} = \mathrm{e}^{w[\log b]}\mathrm{e}^{2k\pi\mathrm{i}w} \quad (k = 0, \pm 1, \pm 2, \cdots)。$$

对 k 的不同值，此表达式给出完全不连接的无穷多个函数。由此，我们得到一个重要的结果：从幂和开方推广而得到的一般的指数表达式 b^w 的值，根本不是一个相互连贯的解析函数，而是无穷多个不同的 w 的函数，其中的每一个都是单值的。

这些函数的各个值肯定是以各种方式相关联的。特别是，当 w 是整数时，它们全部相等，当 w 是分数且其既约式为$\frac{m}{n}$时，它们当中只有有限(n)个彼此不相同者。这 n 个值是 $\mathrm{e}^{\frac{m}{n}\log b} \cdot \mathrm{e}^{2k\pi\mathrm{i}\frac{m}{n}}$，$k = 0, 1, \cdots, n-1$，即所预计的 $\sqrt[n]{b^m}$几个值。

(7) 只有现在，我们才能评论从乘方与开方出发以期达到一个

单值指数函数的传统方法的不当之处。上述方法原来完全像个迷宫，不可能用初等数学方法，特别是在实数范围以内找到出路。如果考虑 b 为负数的情况，在刚得到的结果的启发下，你会清楚地看到这一点。我只是提醒你们，我们现在才能了解对原来看来似乎是任意给出的主值定义 ($b>0$，且 $b^{\frac{m}{n}}>0$) 是合适的。它在我们的无穷多个函数中只给出一个，即

$$[b^w]=e^{w\log b}。$$

另一方面，如果 n 是偶数，则 $b^{\frac{m}{n}}$ 的负实数值将构成一个到处稠密的集合，但属于我们的无穷多个函数中完全不同的一个，且不可能组合而形成一个连续的解析曲线。

现在我想再补充几句，讲讲对数函数的更加深刻的函数论性质。因为当 z 围绕 $z=0$ 旋转一次时，$w=\log z$ 会有一个增量 $2\pi i$，对应的无穷多叶黎曼面在 $z=0$ 必然有一个无穷高阶的支点，使得每绕一圈即从一叶转移到另一叶。如果过渡到黎曼球面，容易看到 $z=\infty$ 是另一个同阶数的支点，此外就再没有支点了。现在就可以弄明白，解某些代数方程时（见第二部分）谈及的对数幂单值化是怎么回事。为了说得确定一些，请考虑一个有理次幂 $z^{\frac{m}{n}}$。借助关系

$$z^{\frac{m}{n}}=e^{\frac{m}{n}\log z},$$

这个幂就成为 $w=\log z$ 的单值函数，这就叫对数幂单值化。为了弄明白这一点，我们同时考虑 $z^{\frac{m}{n}}$ 和对数的黎曼面，两者都展开在 z 平面上。$z^{\frac{m}{n}}$ 的黎曼面有 n 叶，其支点也在 $z=0$ 和 $z=\infty$ 处，在其中每一点，所有 n 叶循环地相连接。如果在 z 平面上设想任一条封闭路线使对数函数沿着此路线回到其原来值（图 7.6），那就意味着在无穷多叶曲面上的路线也是封闭的。很容易看到，当它被映射到 n

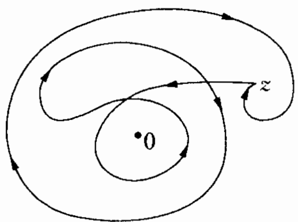


图 7.6

叶曲面时，它的像也是封闭的。从这个几何考虑可推出，当 $\log z$ 回到原来值时，$z^{\frac{m}{n}}$ 总会回到原来值。因此它是 $\log z$ 的单值函数。我很愿作出这个简短解释，因为这里讲的情况是在现代函数论中起重要作用的单值化原理的最简单情况。

现在我们考虑 z 平面及在 z 平面上展布的黎曼面到 w 平面的保角映像，以便把函数关系 $w = \log z$ 的性质弄得更清楚些。为了不致回溯得太远，我们不把对应的球面包括在考虑范围内，尽管包括进去比较好。和前面一样，我们沿实轴将 z 平面分成有阴影（上部）和无阴影（下部）的半平面，其中每一半平面在 w 平面上都有无穷多个像，因为 $\log z$ 是无穷多值的，而由于反函数 $z = e^w$ 是单值的，因此所有这些像必须平滑地相互连接。这意味着，w 平面被许多平行于实轴的线分隔成宽为 π 的平行带域（图 7.7）。这些条带是有阴影和无阴影相间的（实轴上方的第一个带域是有阴影的），它们分别表示上半和下半 z 平面的保角映射，而分隔的平行线对应于实 z 轴。关于详细对应情况，我只指出，仅当 w 保持在一个条带内向左趋向无穷时，z 才总是趋向于零，当 w 向右趋向无穷时，z 变成无穷。反函数 e^w 在 $w = \infty$ 有一个本性奇点。

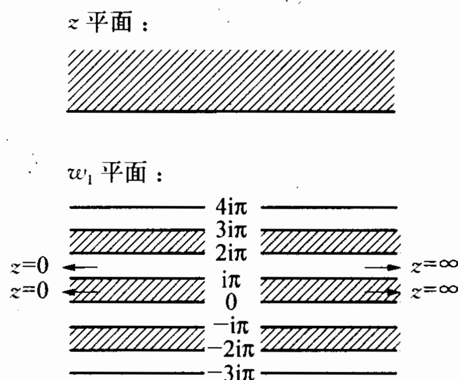


图 7.7

这里绝不要忽略这个表示与皮卡定理之间的联系，因为皮卡定理是新近的函数论中最有趣的定理之一。设 $z(w)$是整超越函数，即只在 $w=\infty$ 处有本性奇点的函数(例如 e^w)。问题为：是否存在这样的 z 值，此函数不可能在有限值 w 处取得 z 值，但当 w 以某种方式趋向无穷时，却是该函数的极限？如果有，又有多少？皮卡定理说，一个函数在一个本性奇点附近最多会遗漏两个值。因此一个整超越函数除去 $z=\infty$ (是必然遗漏的)外，最多还遗漏一个值。e^w 是除∞外只遗漏另一个值 $z=0$ 的函数的一个例子。在划分的每个平行条带内，e^w 趋向于这两个值，但对任何有限值 w，e^w 均不取这两个值。函数 $\sin w$ 是除 $z=\infty$ 外不遗漏其他任何值的例子。

我想再提出曾反复涉及的一点，并应用几何工具加以说明，以结束本节的讨论。我指的就是从幂的极限转变到指数函数的过程，这个过程由公式

$$e^w=\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{nw}$$

给出。如令 $nw=\nu$，则变成

$$e^w=\lim_{\nu\to\infty}\left(1+\frac{w}{\nu}\right)^{\nu}。$$

在取极限前，我们考虑函数

$$f_\nu(w)=\left(1+\frac{w}{\nu}\right)^{\nu},$$

它作为幂的函数论性质是我们已经知道的。它在 $w=-\nu$ 和 $w=\infty$ 各有一个临界点，其底分别为 0 和∞，它将 f_ν 的两个半平面映射成 w 平面上以 $w=-\nu$ 为公共顶点、张角为$\frac{\pi}{\nu}$的一串扇形域(图7.8)。如果 ν 不是整数，则对应于 f_ν 的多值性，这一串扇形域可能覆盖 w 平面有限次或无限次。如果现在令 ν 变成无穷，扇形顶点 $-\nu$ 将无限制向左移，位于$-\nu$ 右边的扇形域就显然变成相应于极限函数 e^w 的

平行条带域,这就对 e^w 的极限定义作了几何解释。通过计算可以证明,在 $w=0$ 的扇形域的张角变成平行带域的宽度 π。

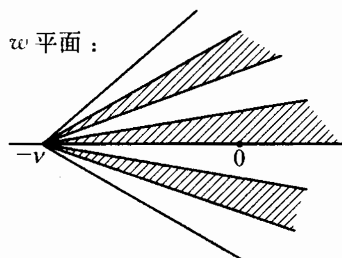


图 7.8

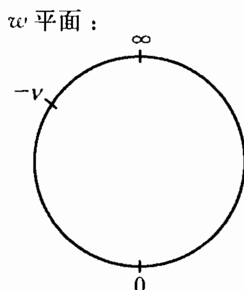


图 7.9

但这里出现一个疑问：如果 ν 连续地变为无穷,不仅通过整数,也通过有理数与无理数,在这时,f_ν 是多值的,将对应于多叶曲面,那么它们是怎样转化到对应于单值函数 e^w 的光滑平面的呢？例如,如果只允许 ν 取分母为 n 的有理数而趋向无穷,则每个 $f_\nu(w)$ 将有 n 叶黎曼曲面。为了了解极限过程,我们暂时考虑 w 球面。对每个 $f_\nu(w)$,w 球由支点在 $-\nu$ 和 ∞ 连接的 n 叶曲面所覆盖。设分支割线是连接此两点的子午线的劣弧段,如图 7.9 所示。如果 ν 趋向 ∞,则支点将重合而分支割线消失。因此,将各叶连接起来的桥梁被摧毁了,出现了 n 个分离的叶,对应它们的是 n 个单值函数,其中有一个是我们的 e^w。如果现在让 ν 通过所有实数,一般将有无穷多叶的曲面,它们的连接处在极限过渡时被破坏了。这些曲面中的每一个曲面,有一叶上的值收敛于分布于光滑球面上的单值函数 e^w,而在其他叶上的值序列一般没有任何极限。这样,对多值幂函数通过极限变成单值指数函数的复杂而奇妙的途径就有了一个完整的解释。

由上面这些考虑可以得出一个普遍的结论,只有把那类问题放到复数域内才有可能对它们获得完整的了解。难道这还不足以成为在中学里教复变函数的理由吗？马克斯·西蒙(Max Simon)就是支

持类似要求的一个人。然而，中等程度的学生，即使在最高年级，我也很难相信能接受这样深的内容。因此我想，在中学里应放弃那种代数分析方法而采用前面已经讲过的简单而自然的方法。不过我还是希望教师充分掌握这里所谈及的所有与函数论的联系，因为教师的知识应大大超过他的讲课范围，他应该熟悉哪里有艰难险阻，才能指导学生安全通过。

经过这一番详细的讨论，在对角函数作相应的考虑时就可以简略一些了。

第八章

角　函　数

在开始之前，我要说明用角函数这个名称似乎比习惯上用的三角函数要好，因为三角学只是这些函数的一个特殊应用。它们本身与指数函数相类似，但其中的反函数又类似对数函数。我们称这些反函数为测圆函数。

§8.1　角函数理论

作为理论考虑的出发点，让我提出一个关于在中学介绍角函数的最好方法的问题。我想最好的方法还是利用求面积的这个总原则。从测量圆弧开始的传统方法，我觉得不十分明显，特别是因为它使初等数学与高等数学之间失去了简单的内在联系。

我再一次直接应用解析几何。在单位圆

$$x^2 + y^2 = 1$$

上考虑由原点到点 $A(x = 1,\ y = 0)$ 和由原点到点 $P(x,\ y)$ 的射线形成的扇形（图 8.1）。为了与常用的符号一致，将用 $\frac{\varphi}{2}$ 来表示这个扇形的面积（因为弧长的习惯符号是 φ）。

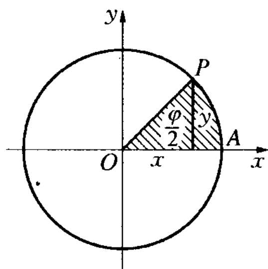


图 8.1

定义 φ 的角函数即其正弦与余弦分别为扇形 $\frac{\varphi}{2}$ 的端点 P 的坐标 x 和 y

$$x = \cos\varphi, \quad y = \sin\varphi。$$

这个符号的起源不清楚。“sinus”(正弦)这个词或许是由阿拉伯词误译成拉丁词的。因为我们没有从弧出发,不能借用习惯的术语 arcsin 和 arccos 来表示反函数,即作为坐标的函数的双倍扇形面积。但根据类比,自然可称$\frac{\varphi}{2}$为正弦(或余弦)的“面积”area,并写成

$$\varphi = 2\text{areasin}\, y = \arcsin y,$$

$$\varphi = 2\text{areacos}\, x = \arccos x。$$

在英美则通用下列记号

$$\varphi = \cos^{-1} x, \quad \varphi = \sin^{-1} y。$$

其他两个角函数

$$\tan\varphi = \frac{\sin\varphi}{\cos\varphi},\ \cot\varphi = \frac{\cos\varphi}{\sin\varphi}$$

(在过去的三角学中还有正割和余割)简单地被定义为两个基本函数的有理组合。它们只是为了简化实际计算而引入的,对我们没有重要的理论意义。

如果随 φ 的增加观察点 P 的坐标,则立即可得到在直角坐标系内定性地表示余弦和正弦的曲线。它们是以 2π 为周期的、人所熟知的波浪形曲线(图 8.2),π 定义为整个单位圆的面积,而不是像通常那样定义为半圆的弧长。

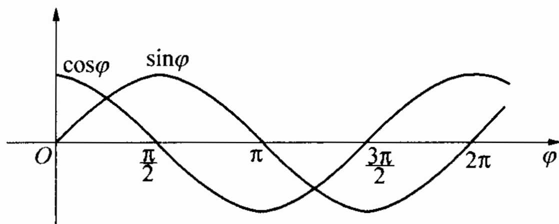


图 8.2

请再把我们引入指数函数与对数函数的方法与这些定义作一比较。你们会记得，我们是从以渐近线为坐标轴的等轴双曲线

$$\xi \cdot \eta = 1$$

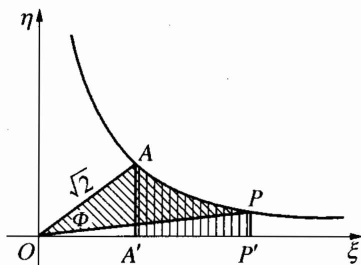


图 8.3

出发的。此双曲线的半轴为 $OA = \sqrt{2}$（图 8.3），而圆的半径为 1。现在我们考虑介于固定坐标 AA' $(\xi = 1)$ 和变动坐标 PP'之间的条带区域的面积。如果称其为 Φ，则有 $\Phi = \log \xi$，且点 P 的坐标可用Φ表达成

$$\xi = e^{\Phi}, \quad \eta = e^{-\Phi}。$$

你会注意到和前面讨论有点类似，但有两个方面是不一样的。首先，Φ 不是前面所说的扇形面积。其次，现在的两坐标是通过一个函数 e^{Φ} 的有理式来表达，而就圆的情况来说，为了获得有理表达式，不得不借助 sin 和 cos 两个函数。不过我们会看到，这种区别很容易得到解决。

首先注意到，三角形 $OP'P$ 的面积为$\frac{1}{2}\xi\eta = \frac{1}{2}$，与点 P 的位置无关，特别是它和三角形 $OA'A$ 的面积相等。因此，如果将后者加到 Φ 内而后从这个和中减去前者，则 Φ 可定义为顶点 A 与变点 P 的两矢径之间的双曲线扇形的面积，完全和圆的情况一样。但仍有符号上的差别。从点 O 看去，在前一种情况下，弧 AP 是逆时针方向，而现在成了顺时针方向。按 OA 将双曲线反射，即变换 ξ 和 η，即可消除这个差别。于是，取点 P 的坐标为

$$\xi = e^{-\Phi}, \quad \eta = e^{\Phi}。$$

最后，将图 8.3（在对 OA 反射后）旋转 45°，使双曲线的主轴为

坐标轴(图 8.4[①])。如果称新坐标为(X, Y),则坐标变换式为

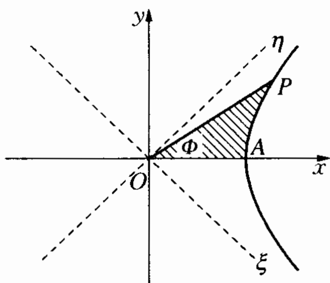


图 8.4

$$X=\frac{\xi+\eta}{\sqrt{2}},\quad Y=\frac{-\xi+\eta}{\sqrt{2}}。$$

双曲线方程于是变成

$$X^2-Y^2=2。$$

现在扇形 Φ 与圆中的扇形$\frac{\Phi}{2}$位置就完全一样了。点 P 的新坐标作为 Φ 的函数可以写成

$$X=\frac{e^{\Phi}+e^{-\Phi}}{\sqrt{2}},\quad Y=\frac{e^{\Phi}-e^{-\Phi}}{\sqrt{2}}。$$

剩下需要将整个图形按 $1:\sqrt{2}$ 的比例缩小,以便像在圆的情况下一样,使双曲线的半轴为 1 而不是$\sqrt{2}$。于是所述的扇形面积和前面的情况完全一样成了$\frac{\Phi}{2}$。如果还是记新的坐标为(x, y),则它们作为 Φ 的函数为

$$x=\frac{e^{\Phi}+e^{-\Phi}}{2},\quad y=\frac{e^{\Phi}-e^{-\Phi}}{2}。$$

它们满足关系式

$$x^2-y^2=1,$$

即双曲线方程。这些函数称为双曲余弦与双曲正弦,并记为

$$x=\cos\mathrm{h}\Phi=\frac{e^{\Phi}+e^{-\Phi}}{2},$$

---

① 译者加。

$$y = \sin h\Phi = \frac{e^{\Phi} - e^{-\Phi}}{2}。$$

最后的结果是：如果用完全一样的方法，分别处理半轴为 1 的圆和等轴双曲线，则一方面得到通常的角函数；另一方面得到双曲函数，因而这些函数彼此完全对应。

你们已经知道，cos h 和 sin h 在许多情况下很有用处。然而，就有关双曲线的处理而论，我们实际退后了一步。首先，坐标(ξ, η)原来可以用一个函数 e^{Φ} 表示，而现在得用两个函数来表示，但它们可用一个代数关系（双曲线方程）联系起来。自然希望像原来对双曲线的推导一样，用反演的方法来处理角函数，只要不反对使用复数，这是十分容易做到的。事实上，如同用 e^{Φ} 表达 $\cos h\Phi$ 与 $\sin h\Phi$ 那样，可以引出一个基本函数将 $\sin\varphi$ 和 $\cos\varphi$ 有理地表达出来。因此这个函数可以在角函数理论中起主要作用。

为此，对圆方程 $x^2 + y^2 = 1$（其中 $x = \cos\varphi$, $y = \sin\varphi$）引入新坐标

$$x - \mathrm{i}y = \xi, \quad x + \mathrm{i}y = \eta,$$

从而有

$$\xi \cdot \eta = 1。$$

现在，正和上述双曲线的情形一样，所希望的中心函数正是第二个坐标 η。如果用 $f(\varphi)$来表示它，则通过变换方程，我们有

$$\eta = f(\varphi) = \cos\varphi + \mathrm{i}\sin\varphi, \quad \xi = \frac{1}{f(\varphi)} = \cos\varphi - \mathrm{i}\sin\varphi。$$

由此得到

$$\cos\varphi = \frac{\xi + \eta}{2} = \frac{f(\varphi) + [f(\varphi)]^{-1}}{2}$$

$$\sin\varphi = \frac{-\xi + \eta}{2\mathrm{i}} = \frac{f(\varphi) - [f(\varphi)]^{-1}}{2\mathrm{i}}。$$

与先前 $\cos h\Phi$，$\sin h\Phi$ 与 e^{Φ} 之间的关系完全类似。如果从一开始就突出圆和双曲线的类比，则欧拉的伟大发现 $f(\varphi)=e^{i\varphi}$ 就没有通常的那种神秘色彩了。

现在提出一个问题：不脱离实数域，能否将 $\cos w$ 和 $\sin w$ 类似地简化为用一个基本函数来表达？如果从射影几何的观点来看我们的图形，那么这一点确实是可能的。事实上，在双曲线情形下，可以视坐标 η 为平行线束 $\eta=$ 常数的参数，此坐标 η 正是所需要的基本函数。从射影观点来看，这意味着有一个顶点在双曲线上（这里是位于无穷远点）的线束。如果在圆和双曲线的两种情况下，都把任何这样线束的参数设想成一个面积的函数，则同样得到一个只用到实数的基本函数。

现在我们来考虑圆和通过点 $S(-1,0)$ 的线束

$$y=\lambda(x+1),$$

其中 λ 为参数（图 8.5①）。在第一部分里，我们曾找到，圆和对应于 λ 的射线的交点 P 的坐标为

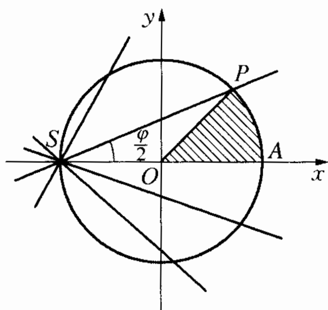


图 8.5

$$x=\cos\varphi=\frac{1-\lambda^2}{1+\lambda^2},$$

$$y=\sin\varphi=\frac{2\lambda}{1+\lambda^2}。$$

所以

$$\lambda=\lambda(\varphi)=\frac{y}{1+x}。$$

事实上是一个合适的实的基本函数。又因为，$\angle PSO=\frac{1}{2}\angle POA$，

① 英译本中角 $\frac{\varphi}{2}$ 的标注位置有误。——编辑

$\angle POA = \varphi$，立即可推出 $\lambda = \tan\frac{\varphi}{2}$。这种通过 $\tan\frac{\varphi}{2}$ 对 $\sin\varphi$ 和 $\cos\varphi$ 作单值表示的方法，在三角学计算中是常常用到的。

在最后一个公式中，可以得到一个关系式

$$\lambda = \frac{y}{1+x} = \frac{1}{\mathrm{i}}\cdot\frac{f-f^{-1}}{f+f^{-1}+2} = \frac{1}{\mathrm{i}}\cdot\frac{f^2-1}{f^2+1+2f}$$
$$= \frac{1}{\mathrm{i}}\cdot\frac{f(\varphi)-1}{f(\varphi)+1}。$$

它表示出 λ 和原先的基本函数之间的关系，其逆则是

$$f(\varphi) = x+\mathrm{i}y = \frac{1-\lambda^2+2\mathrm{i}\lambda}{1+\lambda^2} = \frac{1+\mathrm{i}\lambda}{1-\mathrm{i}\lambda}。$$

因而引入 λ 就相当于确定 $f(\varphi)$ 的一个线性有理函数，使在单位圆圆周上取实值。用这种方法，公式变成实的，但比直接用 $f(\varphi)$ 表示要复杂些。

考虑到用实数更繁一些，是否愿意不用实数而用复数，当然取决于是否会用复数。在这方面值得指出，物理学家早就改用复数，特别是在光学中。例如，在处理振动方程时就是这样。工程师，特别是电气工程师，使用矢量图时，近来也在利用复数，这带来了很大的方便。因而可以说，复数量的使用终于开始普及了，尽管现在大多数的人仍以使用实数为限。

下面简单地说说角函数理论的进一步发展。下一步，我们考虑某些基本规律：

(1) $\sin\varphi$ 的加法定理是

$$\sin(\varphi+\psi) = \sin\varphi\cos\psi+\cos\varphi\sin\psi。$$

对 $\cos(\varphi+\psi)$ 有对应的公式。由于这里所处理的不是真正的初等函数，因此这些公式显得比指数函数的相应公式难。我们的基本函数是 $f(\varphi) = \cos\varphi+\mathrm{i}\sin\varphi$，它满足很简单的关系

$$f(\varphi+\psi)=f(\varphi)\cdot f(\psi)。$$

这和 e^{φ} 的公式完全一样。

(2) 现在很容易求得倍角和半角公式。我只谈其中的两个公式

$$\sin\frac{\varphi}{2}=\sqrt{\frac{1-\cos\varphi}{2}},\quad \cos\frac{\varphi}{2}=\sqrt{\frac{1+\cos\varphi}{2}}。$$

因为它们在建立第一个三角函数表时是很重要的。所有这类关系式可用棣莫弗公式

$$f(n\varphi)=[f(\varphi)]^n,\text{其中 } f(\varphi)=\cos\varphi+\mathrm{i}\sin\varphi$$

很漂亮地给出。棣莫弗是一个法国人，但居住于伦敦，和牛顿有来往，1730 年在他的书 *Miscellanea analytica*（《分析杂记》）内发表此公式。

(3) 从我们的 $y=\sin\varphi$ 的原始定义出发，当然很容易导出反函数 $\varphi=\sin^{-1}y$ 的积分表达式。图 8.6 中所示面积由单位圆的扇形 $\frac{\varphi}{2}(AOP)$ 加上三角形 $OP'P$ 组成，它的边界为坐标轴、坐标 y 处平行于 x 轴的直线和曲线 $x=\sqrt{1-y^2}$。此面积是 $\int_0^y\sqrt{1-y^2}\,\mathrm{d}y$。因三角形的面积为

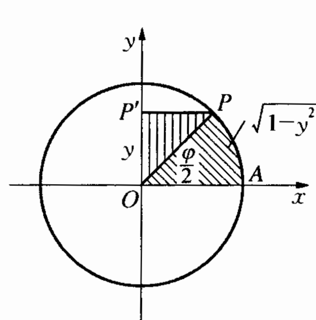


图 8.6

$$\frac{1}{2}OP'\cdot P'P=\frac{1}{2}y\sqrt{1-y^2},$$

故有

$$\int_0^y\sqrt{1-y^2}\,\mathrm{d}y=\frac{1}{2}y\sqrt{1-y^2}+\frac{1}{2}\varphi。$$

由此通过简单的变换可得

$$\varphi=\sin^{-1}y=\int_0^y\frac{\mathrm{d}y}{\sqrt{1-y^2}}。$$

现在可以像对数的情况一样把讨论进行下去，即通过二项式定理展开被积函数，然后仿照默卡多的作法逐项积分。这样可以给出 $\sin^{-1}y$ 的幂级数，由它再作反演，可得到 $\sin\varphi$ 的级数。如同在前面第一部分所看到的，这正是牛顿自己所用的办法。

(4) 然而，我宁愿借用泰勒的伟大发现所提供的简捷方法。根据上面的积分公式，可求得 $\sin\varphi$ 的微分方程如下

$$\frac{\mathrm{d}\sin\varphi}{\mathrm{d}\varphi}=\frac{\mathrm{d}y}{\mathrm{d}\varphi}=\sqrt{1-y^2}=\cos\varphi,$$

由此推出

$$\frac{\mathrm{d}\cos\varphi}{\mathrm{d}\varphi}=-\sin\varphi。$$

由泰勒定理给出

$$\sin\varphi=\frac{\varphi}{1!}-\frac{\varphi^3}{3!}+\frac{\varphi^5}{5!}+\cdots,$$

$$\cos\varphi=1-\frac{\varphi^2}{2!}+\frac{\varphi^4}{4!}-\cdots。$$

很容易看到这些级数对每一个有限值 φ(包括复数)均收敛，因此 $\sin\varphi$ 和 $\cos\varphi$ 是定义在整个复平面上的单值整超越函数。

(5) 如果将此级数与 e^{φ} 的级数作比较，可以看到，基本函数 $f(\varphi)$ 满足关系

$$f(\varphi)=\cos\varphi+\mathrm{i}\sin\varphi=e^{\mathrm{i}\varphi}。$$

这个结果是毫无疑义的，因为 $\sin\varphi$ 和 $\cos\varphi$ 如同 e^{Φ} 一样都是单值整函数。

(6) 剩下的是讨论复函数 $\sin w$ 和 $\cos w$ 的性质。首先注意到，$w=\sin^{-1}z$ 和 $w=\cos^{-1}z$ 中的每一个反函数，都产生一个具有无穷多叶，且在 $+1$，-1，∞ 有交点的黎曼曲面。事实上，在 $z=+1$ 和 $z=-1$ 上有无穷多个一阶的支点，而在 $z=\infty$ 具有两个无穷阶的支点。为了更好地、详尽地找到各叶的走向，我们考虑对应于上(阴

影)下(无阴影)半 z 平面的 w 平面的区域划分。对 $z=\cos w$ 来说,这个划分是由实轴和通过点 $w=0, \pm\pi, \pm 2\pi, \cdots$ 的平行于虚轴的各直线组成,从而成为许多矩形域[①](图 8.7),所有这些区域都延伸到无穷,且有阴影者与无阴影者相间。在点 $w=0, \pm 2\pi, \pm 4\pi\cdots$(对应于 $z=+1$)和点 $w=\pm\pi, \pm 3\pi, \cdots$(对应于 $z=-1$)均各有4个矩形域相遇。这些对应于黎曼曲面的4个半叶,它们在对应的位于 $z=\pm 1$ 之上的支点中的每一个支点处相连接。如果 w 在任何一个矩形域内趋向无穷,$\cos w$ 也趋向于值 $z=\infty$。因而存在着两个分离的无穷个角形域的集合,且每个角形域都延伸至无穷,对应于黎曼面上有两个彼此分离而在 $z=\infty$ 处连接的各无穷多叶的集合。

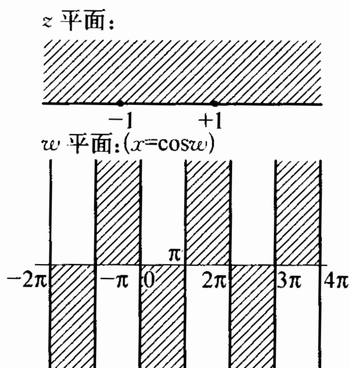


图 8.7

对 $z=\sin w$,除去在 w 平面的图像向右平移 $\frac{\pi}{2}$ 外,其余的均类似。

这些图形,使前一节关于在 $w=\infty$ 处的本性奇点的性质和皮卡定理的关系又得到证实。

§8.2 三角函数表

在简单说明了角函数的理论之后,我想讨论对实际工作极为重要的某些内容,即三角函数表。与此同时,我还要谈及对数表。由于对数表和三角函数表从一开始到现在都是紧密连在一起的,因此,到现在为止,我始终把对数表暂时放在一边。对数表是怎样达到今天这种形式的,这对中学和大学的数学教师都是异常重要和有意义的。

---

① 原文作三角形,是把无穷远点作一个顶点。——中译者

在这里,我当然无法详尽介绍这个表的漫长发展史,但我将引用若干最有意义的著作,努力向你们介绍历史的梗概。至于其他一些著作,其中有的是同等重要的,我建议你们参考特罗夫克的著作。而就对数表而言,可参考密姆克《百科全书》中有关算术计算(numerisches rechnen)的详尽解说,同时参考法文版《百科全书》中道加涅(d'Ocagne)对此所作的修正[①]。这样,历史的叙述就可能完整无缺了。

由于在对数发明之前三角函数表已经发展起来,所以我首先来介绍纯粹的三角函数表。

1. 纯粹的三角函数表

这类表古已有之,其中第一份是所谓“弦表”。

(1) 弦表。此表由托勒密(Ptolemy)为研究天文学的需要而编制,编制时间约在公元 150 年,载于他的著作 *Megale Syntaxis* 中。托勒密在这本著作中建立了以他的名字命名的天文学理论体系(现有一现代版本,海伯格编,1898—1903 年莱比锡版)。此书是由阿拉伯人传给我们的,用 *Almagest* 作书名,这个字可能是由阿拉伯文冠词“al”与残缺的原希腊文书名组合而成。本表按 30 分的间隔列出。它并未直接给出角 α 的正弦,只给出了对应弧的弦长$\left(即\frac{2\sin\alpha}{2}\right)$。这些弦长的值用三位六十进制分数,即 $\frac{a}{60}+\frac{b}{3\,600}+\frac{c}{216\,000}$ 来表示,其中 a, b, c 是 0 与 59 之间的整数。其中的 a, b, c 自然是用希腊数码表示的,即由希腊字母组合而成的,这对我们来说读起来有困难。该表也给出差值,可以按分进行插值。在计算时,托勒密主要用到三角函数加法定理,而后者又表示为一个关于内接四边形的几何

---

① 《数学百科全书》法文版第 1 卷,第 23 页。同时参考 F·凯乔里著《数学名史》,1919 年麦克米伦版,以及 D·E·史密斯著《数学史》,1925 年吉恩出版公司。

定理(托勒密定理)。他也用到前面$\frac{\sin\alpha}{2}$的公式(除有理运算外加上开平方根),还用到内插过程。

(2) 三角函数表在欧洲首次编成,已是1 000多年以后的事了。第一个做出贡献的人是雷乔蒙塔努司(Regiomontanus, 1436—1476年),他的真名是乔安尼·缪勒(Johannes Müller)。他计算了几个三角函数表,从六十进制改成了纯十进制。现在我们都知道正弦、余弦的单位圆线段表示法,但当时还没有人想到这些线段都是分数。三角函数的计算都是利用半径很大的圆来作的,所以这些线段都成了整数了。事实上,这些大数本身都是用十进制写的,但从选择半径可以看到,往往固执地打算采用六十进制。因此雷乔蒙塔努司造第一个表时半径取为6 000 000,直到他造第二个表时,才选了纯粹的十进制10 000 000,并与十进制取得了完全一致。只要简单地插入小数点,就变成今天的十进制表。雷氏的这些三角函数表是在他去世后很久才第一次出版的,附于他的老师G·普尔巴赫(Peurb ach)的作品*Tractatus super propositions Ptolemaei de sinubus et chordis*(纽伦堡,1541年)中。值得注意,这个著作像其他许多数学的基本著作一样,是16世纪40年代印行于纽伦堡的,雷氏本人大半生在纽伦堡度过。

(3) 现在我向你们介绍一本具有最伟大的普遍意义的著作:哥白尼(Copernicus)写的*De revolutionibus orbium coelestium*(《天体运行论》)。此书确立了哥白尼的天文学理论体系。哥白尼从1473年到1543年住在多恩(Thorn),他的著作在雷乔蒙塔努司的三角函数表发表两年后,同样在纽伦堡出版。由于哥白尼从来没有看过这些表,他只好为自己计算小小的正弦表以满足他的理论的需要,此表可在他的书上找到。

(4) 这些表绝对适应不了天文学家的需要,因而哥白尼的一个学生和朋友不久就打算做一件更宏伟的工作。这个人的名字叫赖蒂克斯(Rhäticus),他生于1514年,死于1576年,是魏登堡(Wittenberg)的

教授。你们必须把这一切事情同当时的总的历史背景联系起来看。当时我们处于宗教改革时代，你们知道，魏登堡和自由城市纽伦堡是学术活动的中心。但是在宗教改革斗争中，政治活动和学术活动的重心逐渐地由城市移向君主的宫廷。所以尽管以前一切学术著作印行于纽伦堡，但赖蒂克斯的伟大的三角函数表却是在帕拉丁(Palatine)的赞助下出版的(海德堡，1596 年)，因而冠以 Opus Palatinum 的名字。这个三角函数表是在赖蒂克斯死后不久印行的，比以前的表完善得多，包括有按 10 分间隔的十位三角函数的值。不过说实在的，有许多错误。

(5) 此表后经西里西亚格龙堡的皮蒂斯克斯(Pitiscus，1561—1613 年)的极大改进后出版。皮氏是帕拉丁的教士。这本又是在君主资助下印行的《数学宝书》(法兰克福，1613 年)，包括按 10 分间隔的 15 位三角函数值。同赖蒂克斯的原表比较，消除了许多差错，也更为简明。

我们必须记住，所有这些表主要是依靠半角公式和内插法编成的，因为当时还没有 $\sin x$ 和 $\cos x$ 的无穷级数。所以这些伟大著作中所凝聚的大量劳动和智慧，是应该得到正确的评价的。

2. 对数三角函数表

紧接着这些表之后的是三角对数表的出现。这是异常的巧合，有人可能说是历史的嘲弄。在皮蒂斯克斯把三角函数表改进之后一年，第一个对数表出现了，使这些函数表变得了多余。从此以后，人们用正弦对数和余弦对数来代替正弦和余弦。我已经提到过由纳皮尔造的第一个对数表。

(1) 纳皮尔的 *Mirifici Logarithmorum Canonis Descriptio* 造于 1614 年。他的本意是想简化三角函数的运算，因而他没有给出自然数的对数，而是给出三角函数按一分间隔的七位对数。

(2) 现在的对数表主要由英国人亨利 · 布里格斯(Henry

Briggs，1556—1630 年)造的，他和纳皮尔有来往。他认识到以 10 为底的对数对实数计算有极大优点，因为它更适合于十进制，因此，早在 1617 年，在 *Logarithmorum Chilias Prima* 一书中，他就用这个底代替了纳皮尔用的底，造了以他的名字命名的"人造"或常用对数。为了计算这些对数，布里格斯发明了一系列有趣的方法，使人们能按自己所选择的精确度来确定每一个对数。布里格斯的第二本著作的书名为《对数算术》。在这本书里，他列出了自然数的对数表而不是纳皮尔列的那种角度比的对数表。肯定地讲，布里格斯并未完成计算，他只给出了从 1 到 20 000 和 90 000 到 100 000 之间的整数对数，但有 14 位。值得注意的是，恰巧是旧表给出的位数最多，而现在为了大多数目的，很少几位数就行了。稍后我还会讨论这一点。布里格斯在他的 *Trigonometria Britannica* 中也编出了以分为间隔的三角函数的十位常用对数。

(3) 布里格斯表的空隙，后由数学家、书刊出版销售商、荷兰人艾德里安·弗拉克(Adrian Vlacq)所填满。他为布里格斯表发行了第二版，其中有从 1 到 1 000 000 所有整数的十位对数。我们可以把这本书当作现今所有自然数对数表的依据。

关于对数的进一步发展，在此我只能大致叙述随后的年代中与初期相比所取得的进展的要点。

(a) 第一个本质的进展是在理论方面，即使用对数级数，这是计算对数的极其有用的新方法。早期的表的编者们并不知道这些级数。我们已经看到，纳皮尔计算对数的时候用的是差分方程，即用逐次累加$\frac{\Delta x}{x}$的方法和内插法。布里格斯应用了重要的求平方根的方法，他利用了纳皮尔在其 *Construction* 一书中谈到的结果：知道 a，b 的对数后，即可知道 $\log\sqrt{a\cdot b}=\frac{1}{2}(\log a+\log b)$。或许弗拉克也是用这种方法计算的。

(b) 后来在排印方面作出了改进，用较少的篇幅把更多的材料

收进表中，而且也更醒目。

(c) 最重要的是，通过对旧表的仔细校核，准确性也大大提高，减少了数字的差错，特别是末位数字的差错。

在大量作了改进的表中，我只谈谈最有名的一个——*Thesaurus Logarithmorum Completus*(《对数大全》)。

(4)《对数大全》是奥地利炮兵军官维加(Vega)编的，1794 年出版于莱比锡。此书已极难得，但 1896 年在佛罗伦萨出现了直接照相本。该书包括有自然数和三角函数的十位对数，其编排形式从此成为标准的编排式。例如，在里面有便于插值的小差表。

19 世纪以来，由于在 20 年代把对数部分地引入了中学教学，而且由于在物理和技术上得到了越来越多的应用，因此对数得到了空前的普及。与此同时，位数减少了。因为，就中学和技术上的需要而言，表最好不要太厚，特别是就所需要的精确度而言，三四位数差不多就够了。说实在的，我上学的时候还用七位表，据说理由是这样可以使学生得到"数的尊严"的印象。我们今天的思想一般是比较讲实用的，我们现在到处用二位、三位表，最多用五位表。今天我要给你们看 3 个随意挑出的现代对数表。一个是舒伯特编的袖珍四位表①，你们能在其中发现一切表现手段，如双色套印，每一页上下都重复印上首尾数之类，以排除误解。第二个是亨丁顿(Huntington)编的现代美国对数表②，编排得更为巧妙，表页上有凹凸刻痕，能使你立刻翻到所要找的那一页。最后我要给你们看一把计算尺，正如你们所知道，其实也是一种三位对数表，不过以一种拉尺形式出现，非常方便。你们当然都已熟悉这种工具，现在每一个工程师都随时带在身边。

今天我们还没有达到发展的顶点，但可以很明显地看出进一步

① *Viverstellige Tafeln und Gegetafeln*，现为舒伯特及豪斯纳(Haussner)合编，Sammlung Göschen 出版社，莱比锡，1917 年。

② C·V·亨丁顿，《四位对数表》，节略版，马塞诸塞州剑桥，1907 年。

发展的方向。最近,手摇计算机已进入广泛使用的领域,由于直接相乘起来迅速而可靠得多,所以使对数表已成为多余。不过目前这种机器非常昂贵,只有大办事机构才能购置。当这种机器变得很便宜的时候,数值计算的一个新阶段便宣告开始。就测角术而言,皮蒂斯克斯的老对数表尽管诞生不久就过时了,但到那个时候将会得到应有的地位,因为能够直接提供三角比,使手摇计算机得以立刻运算,从而避免使用对数表。

§8.3 角函数的应用

现在剩下的是给你们概括地介绍一下角函数的应用,我将介绍3个领域。

(a) 三角学。事实上,它为角函数的发明提供了机会。

(b) 力学。其中特别是小振动理论为它的应用提供了广泛的领域。

(c) 周期函数通过三角函数表示。众所周知,它在种种不同的问题中起到重要作用。

让我们立即讨论第一个主题。

1. 三角学,特别是球面三角学

我们在这里面对的是一门非常古老的科学,这门科学在古埃及达到了高度的繁荣,由于两门重要科学的需要而促进了它的发展。当时,大地测量学要求有平面三角的理论,而天文学需要球面三角的理论。在天文史方面,A · v · 布劳恩米尔(A. v. Braunmühl)著 *Vorlesungen über Geschichte der Trigonometrie*[①],是大部头的专著。在三角学的实用方面,最有帮助的是 E · 哈默尔(E. Hammer)的

① 两卷,莱比锡,1900 年及 1903 年。

*Lehrbuch der ebenen und sphärischen Trigonometrie*①；在理论方面，就是我常常提到过的韦伯—韦尔斯坦因的《数学全书》第二卷。

在这些讲稿的限制范围内，我当然不可能系统地谈整个三角学的内容，这属于专门研究的范围。再说我们哥廷根大学设有大地测量学及球面天文学课程，对实用三角学赋予充分的注意。我想向你们谈到的，只是理论三角学中非常有趣的一章。尽管理论三角学这门学问有非常悠久的历史，但还不能认为已没有什么可以研究了，相反还包含着许多尚未解决的问题，而且是相对初级性质的问题，我认为加以研究是会有所获的。我指的是球面三角学。你们会发现这个课题在韦伯—韦尔斯坦因的书中已得到很充分的考虑，该书认为最重要的，是斯特迪（Study）在其奠基作 *Sphärische Trigonometrie, orthogonale Substitutionen und elliptische Funktionen*② 中所发展的思想。我要向你们概括的，是这个范围内的一切理论，并提醒你们注意尚未得到答案的问题。

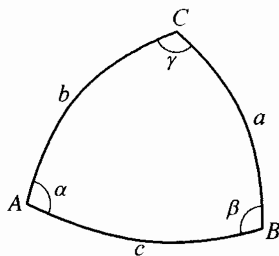


图 8.8

球面三角的基本概念简直不需要解释。在球面上任取 3 点，其中任意两点均非对径点，即可决定唯一的一个三角形。其中每个角和边都在 0 与 π 之间（图 8.8）。进一步的研究表明，最好设边和角都不受限制，可以变到大于 π 或 2π 或 π 的整数倍的情况。于是，必须处理具有重叠的边和绕它的顶点转角多次的情况。此时，对所测量的这些量的符号和量角的方向必须做一些规定。符号原理的重要性之所以得到前后一致的研究，并为这些量在不受限制变化下的一般研究开辟了道路，都要归功于莱比锡的伟大几何学家莫比乌斯（A. F.

① 斯图加特，1906 年（第五版，1923 年）。

② *Abhandlungen der Mathematisch-physikalischen Klasse der Königlich Sächsischen Gesellschaft der Wissenschaften*，第 20 卷，第 2 期，莱比锡，1893 年。

Möbius)。这里，他的研究中具有特别重要意义的部分是 *Entwicklung der Grundrormeln derh sphärischen Trigonometrie in grösstmöglicher Allgemeinheit*[①]。

确定符号的出发点，在于规定绕球面上一点 A 旋转时怎样的角应称为正的(图 8.9)。如果对一点规定好了，就可以确定其他任一点上的正方向，因为从一点能连续地移动到其他点。习惯上从外侧看球时，逆时针旋转为正。其次，必须对球面上的每个大圆指定一个方向。但是我们不能满足于先对某一个大圆确定其方向，再让它连续运动到与第二个大圆重合，进而确定第二个大圆上的方向，因为向两个不同的方向做连续移动，会对第二个大圆得出相反的方向。因此将分别对所考虑的每一个大圆指定一个方向。如果指定的方向不同，则同一个大圆将被看做为不同的图形。据此约定，每一个有向的大圆 a 都可以与一个极点 P 唯一地互相关联，即与其初等意义上的两个极点之一相联系，从这个极点来看，该大圆的方向是正的。反之，球面上每一点有一个有向的极圆。在这些规定之下，三角学里如此重要的极化过程就唯一确定了。

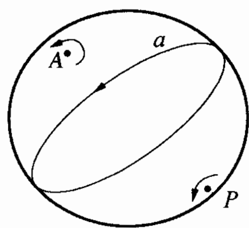


图 8.9

如果在球面上给出了 A, B, C 3 点，还要再作一些规定，才能够使得以这些点为顶点的球面三角形唯一确定。首先，通过 A, B, C 的大圆的方向必须指定，其次还必须知道从点 B 到点 C，从点 C 到点 A 和从点 A 到点 B 转了多少圈。用这种方法确定的长度 a, b, c 可以是任意实数，称之为球面三角形的边。当然，它们都被想象成画在一个半径为 1 的球面上。于是，这些角定义如下：α 是围绕点 A

---

① *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften, mathematisch-physikalische Klasse*，第 22 卷(1860 年)。重印于《莫比乌斯全集》，第 2 卷，第 71 页，莱比锡，1886 年。

使方向CA转到与方向AB一致的正方向的旋转角，可以加上$\pm 2\pi$的任意倍数。另两个角可类似地给以定义。如果现在考察一个普通的初等三角形，如图 8.10 所示，选择各边的方向以使 α, β, γ ①小于 π，则按我们的新定义，角 α, β, γ 是外角，而不是像通常对初等三角形所考虑的内角。

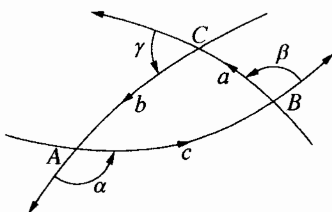


图 8.10

早就已经知道，用这种方法，以补角来代替球面三角形惯用的角，将使球面三角学的公式变得更对称和表达得更为清楚。这样做的深刻理由是基于下面的考虑：由莫比乌斯规定的极化过程，对每一个给定的三角形，唯一地给出了另外一个三角形称为前者的极化三角形。据此新定义，这个极化三角形的边和角，分别就是原来三角形的角和边。于是，根据我们的规定，如果相应地将 a, b, c 改换为 α, β, γ，则每一个球面三角学的公式仍然成立，因而必然总是有这样简单的对称性。另一方面，如果边和角用通常的方法来测量，则这种对称性就丧失了，因为三角形与极化三角形的关系，依赖于人们在给定情况下如何选择边和角以及对未定向的圆如何分辨有疑义的极。

现在已经清楚，在用这种方法定义的一个球面三角形的 6 个元素中，只有 3 个是独立的连续变量，例如两边及其夹角。球面三角形的公式所确定的是这些元素之间的许多关系，或严格些说，是 12 个正弦与余弦之间的代数关系。在这 12 个量中，只有 3 个可以任意改变，其他 9 个则代数地依赖于前者。如果转而使用正弦和余弦，就可以不考虑所增加的 2π 的任意倍数。现在我们把三角学看做是所有可能成立的这种代数关系的总和。因此，按现代的思想方法，可将三角学问题表达如下：如果我们把下列各量

① 原文为 a, b, c。——中译者

$$x_1 = \cos a,\ x_2 = \cos b,\ x_3 = \cos c,$$
$$x_4 = \cos \alpha,\ x_5 = \cos \beta,\ x_6 = \cos \gamma,$$
$$y_1 = \sin a,\ y_2 = \sin b,\ y_3 = \sin c,$$
$$y_4 = \sin \alpha,\ y_5 = \sin \beta,\ y_6 = \sin \gamma$$

视为 12 维空间 R_{12} 的坐标，则对应于真实球面三角形 a, …, γ 的点的集合，形成 R_{12} 内的一个三维的代数图形 M_3，而问题是在 R_{12} 内研究这个 M_3。用这种方法，球面三角就和高维空间的一般解析几何协调一致了。

现在，这个 M_3 必然具有各种简单的对称性。例如极化过程表明，a, b, c 与 α, β, γ 互换总是产生一个球面三角形。用我们的新语言来说，当 x_1, x_2, x_3, y_1, y_2, y_3 分别与 x_4, x_5, x_6, y_4, y_5, y_6 互换时，M_3 内任何一个点变为另一个属于它的点。进而，对应于由 3 个大圆平面将空间分成 8 个卦限，对任何三角形，存在 7 个伴随三角形，它们的各元素是从原来三角形的各元素通过改变符号和增加 π 而得来。这就使得 M_3 的每个点产生 7 个 M_3 的点，它们的坐标 x_1, …, x_6 只是原来的点坐标改变符号而成。这些对称性的集合，构成某个置换群，而改变 R_{12} 的坐标符号，将使 M_3 变换成自身。

现在重要的问题是 M_3 内的坐标所满足的，并组成全部三角学公式的代数方程。因为 $\sin^2\alpha + \cos^2\alpha = 1$，所以一开头，我们有 6 个平方关系

$$x_i^2 + y_i^2 = 1 \quad (i = 1,\ 2,\ \cdots,\ 6), \tag{1}$$

或用几何语言来说，有 6 个二阶圆柱面 $F^{(2)}$ 通过 M_3。

球面三角学的余弦定理给出了另外 6 个公式。用我们的符号，这个余弦定理记作

$$\cos a = \cos b \cos c - \sin b \sin c \cos \alpha。$$

通过极化可得

$$\cos\alpha = \cos\beta\cos\gamma - \sin\beta\sin\gamma\sin a。$$

这些方程,加上通过循环置换 a, b, c 与 α, β, γ 而得到的其他 4 个方程,决定了通过 M_3 的 6 个三次曲面 $F^{(3)}$:

$$\begin{cases} x_1 = x_2x_3 - y_2y_3x_4, \\ x_2 = x_3x_1 - y_3y_1x_5, \\ x_3 = x_1x_2 - y_1y_2x_6, \end{cases} \tag{2}$$

$$\begin{cases} x_4 = x_5x_6 - y_5y_6x_1, \\ x_5 = x_6x_4 - y_6y_4x_2, \\ x_6 = x_4x_5 - y_4y_5x_3。 \end{cases} \tag{3}$$

最后,我们可以利用矩阵

$$\begin{vmatrix} \sin a & \sin b & \sin c \\ \sin\alpha & \sin\beta & \sin\gamma \end{vmatrix} = \begin{vmatrix} y_1 & y_2 & y_3 \\ y_4 & y_5 & y_6 \end{vmatrix}$$

的子行列式等于零来表达的正弦定理,或者把它们全写出来就是

$$y_2y_6 - y_3y_5 = y_3y_4 - y_1y_6 = y_1y_5 - y_2y_4 = 0。 \tag{4}$$

这些表达式代表 3 个二次曲面 $F^{(2)}$,其中实际上只有两个是独立的。因此,对 R_{12} 内的 M_3,一共建立了 15 个方程。

一般来说, $12-3=9$ 个方程怎样也不足以在 R_{12} 内确定一个三维代数图形。甚至在 R_{12} 的普通几何学里,也并非每个空间曲线都能成为两个代数曲面的完全交。最简单的例子是三次代数曲线,它至少要求 3 个方程来确定。很容易看到,在我们的情况下,9 个方程(1)和方程(2)并不能确定 M_3。众所周知,由余弦定理推出正弦定理差一个符号未定,而它通常要通过几何考虑才可确定。然后,我们很想知道:在上述三角学方程中,究竟有多少个方程和哪一些方程可以使 M_3 完全确定?就此而论,我想提出 4 个到现在为止尚未得到准确答案的问题。这些问题也许值得彻底研究。掌握了处理球面三角学的公式之后,研究起来可能并不特别困难。这 4 个问题是:

(1) M_3 的阶数是多少?

(2) 可以把 M_3 完全表示出来的最低次方程是什么?

(3) 表示 M_3 的线性独立方程一共有多少? 也就是: 每个通过 M_3 的曲面均能被写成形式

$$m_1 f_1 + \cdots + m_n f_n = 0$$

时,其中 m_1, …, m_n 是整数,则这样的方程式 $f_1 = 0$, …, $f_n = 0$ 有多少个? 这里所需要的方程,可能比第二个问题中所说的最低次方程多。

(4) 在 n 个表达式 f_1, …, f_n 之间存在着什么代数同一性(所谓链)?

过去已经有人在完全相同的方向上进行过研究,不过问题的提法有些不同,查阅这些研究资料,就可以熟悉以上 4 个问题。1894 年哥廷根大学奇泽姆(G. Chisholm)小姐(现在是杨夫人)的博士论文中,有这方面的研究。顺便说一下,奇泽姆小姐是在普鲁士通过博士学位考试的第一位妇女。奇泽姆小姐的各种初步假设中,最有价值的是选择半角的余切和边作为独立坐标。因为 $\tan\frac{\alpha}{2}$(当然$\cot\frac{\alpha}{2}$也一样)是能将 $\sin\alpha$ 和 $\cos\alpha$ 唯一地表示出来的基本函数,所以可以把所有三角学方程用 $\cot\frac{a}{2}$, …, $\cot\frac{\gamma}{2}$之间的代数方程写出来。于是,在以 $\cot\frac{a}{2}$, …, $\cot\frac{c}{2}$, $\cot\frac{\alpha}{2}$, …, $\cot\frac{\gamma}{2}$为坐标的 6 维空间 R_6 内,球面三角形组成一个三维图形。奇泽姆小姐还指出,M_3 是八阶的,且能作为 R_6 内 3 个二次曲面(二次方程)的完全交而完满地表示出来。她还研究了由此提出的、类似前述的那些问题。

在我的超几何函数的讲义中①,我把前面讨论过的,把边和角的正弦与余弦联系在一起的球面三角公式组称为第一类公式,以便和

① 1893—1894 年冬季学期,E·利特(E. Ritter)整理,1906 年,莱比锡重印。

本质不同的第二类公式区分开来。后者是关于边和半角的三角函数之间的代数方程式，在研究时，最好选用

$$\cos\frac{a}{2},\ \sin\frac{a}{2},\ \cdots;\quad \cos\frac{\alpha}{2},\ \sin\frac{\alpha}{2},\ \cdots$$

这 12 个量作为新的 12 维空间 R'_{12} 的坐标，在 R'_{12} 内，球面三角形仍然组成三维图形 M'_3。在 19 世纪初期，迪兰拜(Delambre, 1807 年)，摩尔维迪(Mollweide, 1808 年)和高斯(1809 年)[①]3 人几乎同时独立发表的那些漂亮的公式，都出自这里。它们就是由以下公式

$$\frac{\sin\frac{\beta+\gamma}{2}}{\sin\frac{\alpha}{2}}=\pm\frac{\cos\frac{b-c}{2}}{\cos\frac{a}{2}},\quad \frac{\sin\frac{\beta-\gamma}{2}}{\sin\frac{\alpha}{2}}=\mp\frac{\sin\frac{b-c}{2}}{\sin\frac{a}{2}},$$

$$\frac{\cos\frac{\beta+\gamma}{2}}{\cos\frac{\alpha}{2}}=\mp\frac{\cos\frac{b+c}{2}}{\cos\frac{a}{2}},\quad \frac{\cos\frac{\beta-\gamma}{2}}{\cos\frac{\alpha}{2}}=\pm\frac{\sin\frac{b+c}{2}}{\sin\frac{a}{2}}$$

作轮换排列而得的 12 个公式。这些公式的本质和新的东西是：与第一类公式相反，这里有双重符号；对同一个三角形，在这些公式中有时同取上面的符号有时同取下面的符号，所得公式均同时成立，而且存在适合两类符号的三角形。换句话说，上述 R'_{12} 内的球面三角形的 M'_3，满足两个完全不同的 12 个三次方程的方程组，从而分成两个分离的代数图形，即：使一种符号的公式成立的图形 $\overline{M}_3$ 和使另一种符号的公式成立的图形 $\overline{\overline{M}}_3$。由于这个重要的事实，这些公式在球面三角学理论中有着十分重要的意义。它们远非旧公式的变换可比，因为旧公式最多只能用于简化三角计算。肯定地说，迪兰拜和摩

① 高斯公式发表于 *Theoria motus corporum coelestium* 第 54 期，1906 年莱比锡重印，收入其论文集第 7 卷，第 67 页。

尔维迪就是只从简化三角计算这个实用观点来考虑这些公式的。只有高斯具有更深的洞察力,他说“如果掌握球面三角思想的最大一般性”,就可以看到符号是可以改变的。因此,我们认为这些公式应以高斯的名字命名,虽然不是他最先发表的。

斯特迪第一个认识到球面三角公式所涉及的全部问题并在1893年的专题研究中进行了深入的探索。要把他的主要结果叙述出来,最方便的办法是考虑以不加限制的6个变量 $a, b, c, \alpha, \beta, \gamma$ 为坐标组成的六维空间,我称这些变量为三角形的超越部分以区别于代数部分 $\cos a$, …或 $\cos\frac{a}{2}$, …,因为前者是三角形顶点的普通空间坐标的超越函数,而后者是代数函数。在这个 R_6 里,所有球面三角形的集合组成超越图形 $M_3^{(t)}$,它在 R'_{12} 内的像是上面考虑过的代数图形 M'_3。因为后者分裂成两部分且映射函数 $\cos\frac{a}{2}$,…都是超越坐标的单值连续函数,所以超越图形 $M_3^{(t)}$ 至少也分裂成两部分。斯特迪的定理如下:由最一般类型的球面三角形的量 $a, b, c, \alpha, \beta, \gamma$ 组成的超越图形 $M_3^{(t)}$,对应于高斯公式的两组符号而分成分离的两部分,每一部分都是一个连通的连续统。这里本质的东西是排除了更进一步的分割。利用三角学公式作进一步的处理,也不可能对球面三角形进行类似的、同等重要的分组。对应于高斯公式中取上面符号的那部分三角形,称为正常三角形,另一部分称为反常三角形。因此,可将斯特迪定理简述为:球面三角形的全体可分解为一个正常三角形和一个反常三角形的连续统。你们可以在韦伯—韦尔斯坦因的书(1903年第二版,第2卷,第385页)中了解到进一步的细节以及这个定理的证明,我在这里只想把结果叙述清楚。

现在必须进一步谈谈两类三角形的差异。如给出一个球面三角形,即给出一组可容许的数组 $a, b, c, \alpha, \beta, \gamma$,它们的正弦与余弦满足第一类方程,从而表示 $M_3^{(t)}$ 中的一点,怎样决定此三角形是正常的还是反常的呢?为回答此问题,我们首先求出所给的数对模 2π 的

最小正同余数 a_0, b_0, c_0, α_0, β_0, γ_0

$$a_0 \equiv a (\bmod 2\pi), \cdots, \alpha_0 \equiv \alpha (\bmod 2\pi), \cdots,$$

$$0 \leqslant a_0 < 2\pi, \cdots, \quad 0 \leqslant \alpha_0 < 2\pi, \cdots。$$

它们的正弦与余弦和 a, …, α, …的相同,所以它们也组成一个三角形,称之为对应于所给三角形的简化或莫比乌斯三角形,因为莫比乌斯不考虑超过 2π 的部分。然后,可以通过一个表来确定莫比乌斯三角形是正常态还是反常态。你们在韦伯—韦尔斯坦因的书(第 352,379,380 页)中可以找到这一点,并找到常态及反常态三角形图(第 348,349 页),不过不太清晰。和通常一样,称在 π 和 2π 之间的角为凹角,为简单起见,也把这个称呼用于球面三角形的边。于是,每类三角形共有 4 种典型情况。

Ⅰ. 正常莫比乌斯三角形

(1) 0 个边是凹的;0 个角是凹的。

(2) 1 个边是凹的;2 个邻角是凹的。

(3) 2 个边是凹的;1 个夹角是凹的。

(4) 3 个边是凹的;3 个角是凹的。

Ⅱ. 反常莫比乌斯三角形

(1) 0 个边是凹的;3 个角是凹的。

(2) 1 个边是凹的;1 个对角是凹的。

(3) 2 个边是凹的;2 个对角是凹的。

(4) 3 个边是凹的;0 个角是凹的。

除上述之外,没有别的情况。因此,这个表使我们能实际确定一个莫比乌斯三角形的特性。

在上述说明之后,按公式

$$a = a_0 + n_1 \cdot 2\pi, \; b = b_0 + n_2 \cdot 2\pi, \; c = c_0 + n_3 \cdot 2\pi,$$

$$\alpha = \alpha_0 + \gamma_1 \cdot 2\pi, \; \beta = \beta_0 + \gamma_2 \cdot 2\pi, \; \gamma = \gamma_0 + \gamma_3 \cdot 2\pi,$$

从对应的简化三角形转移到一般三角形 $a, \cdots, \alpha, \cdots$，然后可以使用下面定理：按 6 个整数的和 $n_1+n_2+n_3+\nu_1+\nu_2+\nu_3$ 是偶数或奇数，相应地决定一般三角形与其简化三角形具有相同或相反的特性。因此，每个三角形是正常或反常的特性就可以确定了。

我将用关于球面三角形的区域的一些说明来结束本章。在斯特迪和韦伯—韦尔斯坦因的书中一点也没有提到这一点，我在所著《球面三角形早期函数论研究》中确实考虑到了。到现在为止，只是把三角形当作满足正弦与余弦定理的 3 个角与 3 条边的一个集合。在我的研究中，我关心到由这些边包围的一个确定的面积，在某种意义上说，关心在这些边之间绷紧的一块膜，还涉及某些角。

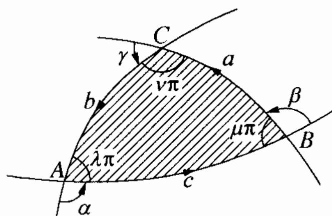


图 8.11

当然，现在不再像以前为了对称的原因而把 α, β, γ 想象为三角形的外角。我们宁愿讨论膜片在顶点所形成的那些角，我称它们为三角形的内角。我将按我的习惯用 $\lambda\pi, \mu\pi, \nu\pi$ 来表示它们（图 8.11）。这些角也可以看成是不受限制的正的变量，因为这个膜可以包着顶点若干次。据此，将用 $l\pi, m\pi, n\pi$ 表示边的绝对长度，它们也是不受限制的正的变量。但边和角不可能再相互独立地“包着顶点”了，即如以前那样包含 2π 的任意倍数了，因为一个具有这种边和角的单连通的膜如果存在，这个事实必定表现为这些包着顶点的次数之间具有某些关系。在专题论文“Überdie Nullstellen der hypergeometrischen Reihe”[①]中，我称这些关系为球面三角形的互补关系。如果用 $E(x)$ 表示小于 x 的最大正整数 $(E(x)<x)$，那么这些关系就是

① 《数学年刊》第 37 卷（1888 年），重印后收于我的《数学论文集》第 2 卷（1921 年），第 550 页。

$$E\left(\frac{l}{2}\right)=E\left(\frac{\lambda-\mu-\nu+1}{2}\right),$$

$$E\left(\frac{m}{2}\right)=E\left(\frac{-\lambda+\mu-\nu+1}{2}\right),$$

$$E\left(\frac{n}{2}\right)=E\left(\frac{-\lambda-\mu+\nu+1}{2}\right)。$$

因为 $E\left(\frac{l}{2}\right)$ 给出了包含在边 $l\pi$ 内的 2π 的倍数，所以知道了角 $\lambda\pi$，$\mu\pi$，$\nu\pi$ 和它们包着顶点的次数之后，根据上面的关系，就可以准确地确定边 $l\pi$，$m\pi$，$n\pi$ 所要求的“包着顶点的次数”。特别是很容易看到，3 个数 $\lambda-\mu-\nu$，$-\lambda+\mu-\nu$，$-\lambda-\mu+\nu$ 之中最多有一个是正数。于是，右边 3 个自变量中只有一个可能超过 1。又因当 $x\leqslant 1$ 时 $E(x)=0$，故只有一个包着顶点的次数可能异于零。换句话说，三角形膜中最多只有一个边会包着顶点（大于 α），而此边必然是最大角的对边。

为了证明这些互补的关系，请你们参考我的油印讲义 *Über die hypergeometrische Funktion*（即前面提过的《超几何函数》——中译者），不过这份讲义早已用完了。在那份讲义中，以及在《数学年刊》第 37 卷的专题论文中，我最初的假设比现在的假设要广一些，考虑了球面上以任意圆而不一定以大圆为界的球面三角形。

下面简单讲一下证明的思路。先假定有一个基本三角形，肯定能在它的上面绷一个膜片。在边上或在顶点上（带有支点）不断地接上圆形膜，逐步得到一个最一般的、可能的三角形膜。图 8.12 用球极投影表示出一个三角形 ABC 为例。它由一个基本三角形与以大圆 AB 为边界的半球面连接而成，从而边 AB 和角 C 都包着顶点。显然，互补关系在此仍然成立。用同样方法可看到，对于能用这种方法建立起来的最一般的三

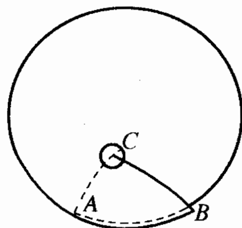


图 8.12

角形膜，互补关系仍然成立。

现在必须问，这些满足互补关系的三角形，怎么会符合我们曾讨论过的一般理论？它们显然是一些特例，以能绷上膜片的三角形为限（因为在一般情况下，边和角的包着顶点的次数是完全任意的）。开头大家可能迷惑不解，因为我们已经看到所有正常三角形（其中某些不一定满足互补关系）组成一个连续统，因此，任何一个正常三角形都可以从一个基本三角形通过连续变形而得出。大家自然会想，在这个变形过程中，绷在最初的基本三角形上的膜不可能失去。如果我们约定，按三角形的边界走向是正方向（逆时针）或反方向而相应地称其区域为正或负，这样将莫比乌斯的符号改变原则推广到区域，这一疑难就释然了。相应地，当一个自交的曲线围成几个局部区域时，整个面积就成了这几个部分面积的代数和，每部分面积的符号按边界走向的正负而定。图 8.13 是带不同阴影部分的差，而图 8.14 则是各部分之和。当然，这些规定仅仅是分析定义本身所提供的几何解释。

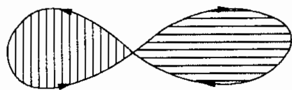


图 8.13

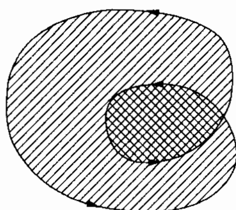


图 8.14

如果将这些讨论应用于由圆弧构成的三角形，则可将每个正常三角形与球面上的一个区域联系起来，使得绕三角形一周时，这个区域的不同部分将会带有不同的符号，因为每个部分的边界将有不同的走向。于是，互补关系成立的那些三角形仅仅是特例，即其区域由一个以正回路为界的单个膜片组成。正是这个特性使三角形研究对于我的早年的研究具有很大的函数论的价值。

我现在用一个例子来说明这个情况。我们考虑三角形 ABC 的

球极投影(图 8.15),其中大圆 BA, CA 交于点 A, A', A 是离弧 BC 较远者。按一般定义,三角形的角为外角。如果现在将定义变为它们的补角(即内角),就会发现 $\mu\pi$ 和 $\nu\pi$ 分别度量了 BC 到 BA 和 CA 到 CB 的旋转角,因此在我们的情况下是正的。类似地,$\lambda\pi$ 度量了 AB 到 AC 的旋转,因此是负的。令 $\lambda=-\lambda'$, $\lambda'>0$, 则三角形 $A'BC$ 显然是具有角 $\lambda'\pi$, $\mu\pi$, $\nu\pi$ 的基本三角形,三个角都是正的。如果现在选择三角形 ABC 的一个回路,则将沿正的方向走过基本三角形 $A'BC$,但沿负的方向走过球面扇形 AA',而三角形 ABC 的面积据莫比乌斯意义就是这两面积之差。设想在点 A' 把膜翻一个面,使得扇形的后侧或负侧被转成前侧,这样也许就能形象化地理解三角形的膜分成正负两部分的情况。据此模式,不难造出更复杂的例子。

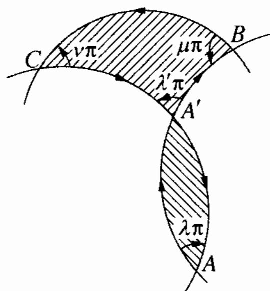


图 8.15

我现在想通过同一个例子来说明,对面积的这种一般定义,基本三角形的面积公式仍然成立。如你们所知,一个单位球面上具有角 $\lambda\pi$, $\mu\pi$, $\nu\pi$ 的球面三角形的面积,由所谓球面角盈余 $(\lambda+\mu+\nu-1)\pi$ 给出,其中 λ, μ, $\nu>0$。我们会看到此公式对上面的三角形 ABC 也成立。显然,基本三角形 $A'BC$ 的面积为 $(\lambda'+\mu+\nu-1)\pi$。还应当把角为 $\lambda'\pi$ 的扇形 AA' 的面积从中减去。因为扇形面积与它的角成比例且当角为 2π 时是 4π(整个球面),故扇形 AA' 的面积为 $2\lambda'\pi$。于是,我们得到 ABC 的面积为

$$(\lambda'+\mu+\nu-1)\pi-2\lambda'\pi=(-\lambda'+\mu+\nu-1)\pi$$
$$=(\lambda+\mu+\nu-1)\pi。$$

如果有一个具有任意边和任意角的正常三角形,套上由若干部分组成的一块膜片,欲确定其面积(按符号规律,为各部分面积的代数和),则其结果将表明公式 $(\lambda+\mu+\nu-1)\pi$ 仍成立,当然这里的

$\lambda\pi$, …是膜片的实际角而不是前面所说的外角。但是,这个研究工作尚未完成。它并不十分困难,我很希望有人担当起来。与此同时,从这一观点去讨论反常三角形所起的作用,也是很重要的。

我将在此结束三角学这个主题而进入角函数的第二个重要应用,它也属于中学教学范围内。

2. 小振动理论,特别是单摆的理论

我将按大学的习惯,用无穷小分析扼要地复习一下摆的定律和推导。一个质量为 m 用长 l 的线悬挂着的摆(图8.16),它偏离平衡位置的角为 φ。由于垂直向下的重力作用,从力学基本定律可推出摆的运动由方程

$$\frac{\mathrm{d}^2\varphi}{\mathrm{d}t^2}=-\frac{g}{l}\sin\varphi \qquad (5)$$

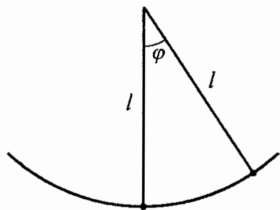


图 8.16

决定。对小振幅,用 φ 取代 $\sin\varphi$ 不会有严重误差。这就给出所谓摆的无穷小振动

$$\frac{\mathrm{d}^2\varphi}{\mathrm{d}t^2}=-\frac{g}{l}\varphi。 \qquad (6)$$

如你们所知,此微分方程的通积分是角函数,如前所述,正是由于它们的微分性质,故在此特别重要。通积分为

$$\varphi=A\sin\sqrt{\frac{g}{l}}t+B\cos\sqrt{\frac{g}{l}}t,$$

其中 A, B 是任意常数。如果引入适当的新常数 C, t_0,则可得

$$\varphi=C\cdot\cos\sqrt{\frac{g}{l}}(t-t_0), \qquad (7)$$

其中 C 称为振幅,t_0 称为初相。由此可得周期 $T=2\pi\sqrt{\frac{l}{g}}$。

这些想法都是十分简单和清楚的,如果想进一步深入,当然可以

用作图方式表示出来。但是，这些道理和在中学里广泛使用的所谓单摆定律的初等处理是多么的不同。其不同在于中学里尽量避免使用无穷小演算，可是恰好是在这里问题的本质需要用无穷小演算。因此，在中学里用的是一种事后硬凑出来的方法，即用到无穷小的概念而不用它们的正确名字。如果要使一切都严格进行，这样的方案当然是极端复杂的，因而处理的方法往往极不完善，稍加思索就看出不能作为单摆定律的证明。于是就出现了一个古怪的现象：同一个教师，当他在教数学课时，对所有结论的逻辑严格性提出了十分高的要求。他在作判断时一心遵守着 18 世纪的传统，他的要求是无穷小计算所不能满足的。但在下一堂课教物理学时，他又接受了很多成问题的结论，并用最大的胆量去使用无穷小。

为了弄得更清楚些，让我来简单叙述一下在教科书里实际讲的，并在教学中使用的单摆定律初等推导的思路。一般从正则摆出发，即一个端点以匀速 v 绕垂直轴做圆周运动使摆的悬线描出一个圆锥面(图 8.17)。在力学中，这叫正则进动。当然，这种运动的可能性在学校里是被假设为已知的，是一种经验，而问题仅限于求出速度 v 和摆的定偏角 $\varphi=\alpha$(摆线描出的圆锥面顶的张角)之间的关系。

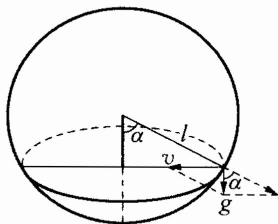


图 8.17

首先注意到，摆头描出一个半径 $r=l\sin\alpha$ 的圆，当 α 充分小时，可以写成 $r=l\cdot\alpha$。然后论及离心力。质量为 m、速度为 v 的摆头必然具有离心力

$$m\frac{v^2}{r}=m\frac{v^2}{l\alpha}。$$

为了保持此运动，必须有指向圆周中心的相等的向心力。于是将重力分解为两个分力：一个力与摆线方向相同；另一个所需要的力在圆周平面上指向圆心，大小为 $m\cdot g\cdot\tan\alpha$ (图 8.17)。由于 α 充分

小，故也可代之以 $mg\cdot\alpha$。于是得到所求的关系式为

$$m\cdot\frac{v^2}{l\alpha}=m\cdot g\cdot\alpha,\text{或 } v=\alpha\sqrt{gl}。$$

这个摆的摆动周期，即走完整个圆周 $2\pi r=2\pi l\alpha$ 的时间为

$$T=\frac{2\pi l\alpha}{v}=2\pi\sqrt{\frac{l}{g}}。$$

换句话说，当摆角 α 充分小时，正则摆形成一个在时间上与 α 无关的正则进动。

现在为了扼要地评论这一部分推导。我们可以允许用 α 取代 $\sin\alpha$和 $\tan\alpha$。我们在上面作准确的推导时也是这样做的，因为这才允许从"有限"振动转到"无穷小"振动。但必须注意到，上面所用到的离心力的公式，只有把所有微小量忽略不计时才能用初等方法推出，而严格论证这一点，只有用微分学才行。例如，离心力的严格定义事实上需要二阶导数的概念，所以即使用初等推导也必须把这个概念偷偷搬进来。既然如此，老师也就讲不清所讲的内容，造成理解上的极大障碍。可是这个障碍在使用了微分学以后就完全不存在了。在此不必详谈，让我继续讨论单摆振动吧。

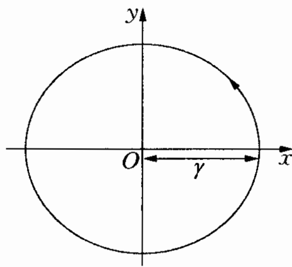


图 8.18

上面的考虑指出了在圆上做匀速运动的可能性。如果在这个圆的平面上（从近似观点来看，即球的切平面）建立坐标系 xy（图 8.18），则用分析力学的话来说，这个运动的方程将是

$$\begin{cases}x=l\cdot\alpha\cdot\cos\sqrt{\dfrac{g}{l}}(t-t_0),\\[2ex] y=l\cdot\alpha\cdot\sin\sqrt{\dfrac{g}{l}}(t-t_0)。\end{cases}\tag{8}$$

但我们希望摆作平面振动，即希望在 xy 平面的摆头沿直线即沿 x 轴运动。这个运动方程必须是

$$x = l \cdot C\cos\sqrt{\frac{g}{l}}(t-t_0),\ y=0。\tag{9}$$

当 $\varphi=\frac{x}{l}$ 时，即可推出正确的方程(3)。因此，必须在不使用动力学微分方程的情况下，从方程(8)得出方程(9)。通过微小振动的叠加原理，有可能做到这一点。叠加原理指出，当运动 x，y 和 x_1，y_1 给出后，运动 $x+x_1$，$y+y_1$ 是可能的。我们可以将逆时针的摆运动(8)和顺时针摆的运动：

$$x_1 = l \cdot \alpha\cos\sqrt{\frac{g}{l}}(t-t_0),\quad y_1 = -l \cdot \alpha\sin\sqrt{\frac{g}{l}}(t-t_0)$$

组合起来。然后如果令 $\alpha=\frac{C}{2}$，则运动 $x+x_1$，$y+y_1$ 就正是所要求的振动(9)。

在评论上面所讲的道理时，首先会问：不用微分学，怎么建立叠加原理？至少会问：怎样使它成为可信？用这些初等方法，逐次忽略微小量，即使每次是许可的，最终是否会累积成显著的误差？对此始终还是有疑问的。我们对此不必详加讨论，因为这些问题完全是初等问题，只要喜欢想，谁都能够想清楚。作为结束语，让我强调一下，在整个讨论中，我们所关心的中心问题是教学问题。首先，显然需要考虑无穷小演算。而且，作为这类一般应用的预备知识，显然需要一个关于角函数的一般性引论，而不只是讲三角形的几何学。

现在来谈角函数的最后一项应用。

3. 用角函数的级数(三角级数)表达周期函数

如你们所知，在天文、数学物理等学科中都经常需要考虑周期函数并用其进行计算。本段标题所表示的方法是最重要、也是最常用到

的。为方便起见，我们将这样选择单位，使给定的函数 $y=f(x)$ 以 2π 为周期(图 8.19)，于是提出这样的问题：能否用带有适当选择的常数因子 x 的 1 倍、2 倍，直到 n 倍等整数倍的余弦与正弦之和，来逼近这个函数？换句话说，能否在一个充分小的误差范围内，用形式为

$$S_n(x)=\frac{a_0}{2}+a_1\cos x+a_2\cos 2x+\cdots+a_n\cos nx+b_1\sin x$$
$$+b_2\sin 2x+\cdots+b_n\sin nx \tag{10}$$

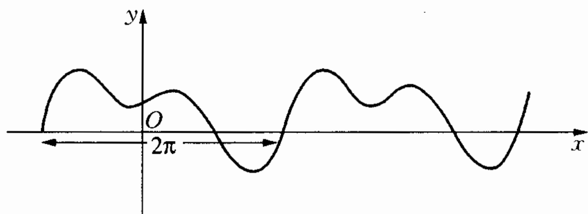


图 8.19

的表达式代替 $f(x)$？在常数项上加上$\frac{1}{2}$因子的目的是为了能对系数给出一个通用的表达式。

首先，我要抱怨教科书里的讲法，这一次所抱怨的是微积分部分。人们不首先提出上面所概述的初等问题，而似乎常常认为，唯一有意义的问题是与我们所提出问题相联系的理论问题，即 $f(x)$是否能严格地用无穷级数来表达的问题。龙格在"Theorie und Praxis der Reihen"①中的讲法，是明显的例外。事实上，就实用而言，理论本身是引不起人们兴趣的，因为实践中只关心有限的几项，项数不会太多。而且理论结果甚至在事后也不能对此级数的实际可用性得出一个结论。从级数的收敛性绝不能得出结论说，前面若干项就给出了与和的相当好的近似。相反，发散级数的前若干项在某些情况下

① *Sammlung Schubert*，第 32 期，莱比锡，1904 年。并可参阅 W · E · 拜尔利(W. E. Byerly)"傅里叶级数及球面调和"一文。

却可能在表示一个函数方面是有用的。我之所以强调这些事,是因为如果有人只知道通常的表达式,而希望把有限项的三角级数用于例如物理实验室,他往往会上当而得不到满意的结果。

有限三角和这一问题早已获得完善的处理,因此对它的习惯性的忽视似乎更值得引起重视。权威性的处理是贝塞尔(Bessel)于 1815 年作出的。关于这些问题的历史及文献,可以查阅《百科全书》中伯克哈特关于三角内插法的详细解说(见《百科全书》Ⅱ A9,第 624 页及随后部分)。此外,这里我们所关心的公式,本质上与一般收敛性证明中的公式相同,只是其中的思想具有其他的细微含义,使材料更合乎实用。

现在详细讨论我们的问题。我首先将提出这样的问题:在给定项数 n 后,怎样最适当地确定系数 a, b, …? 贝塞尔就此发展了最小二乘法的思想。对一个特殊值 x,当我们用三角级数前 $2n+1$ 项的和 $S_n(x)$代替 $f(x)$时,其误差为 $f(x)-S_n(x)$,而代表在整个区间 $0\leqslant x\leqslant 2\pi$ 内的接近程度的度量应是所有误差的平方和,即积分

$$J=\int_0^{2\pi}[f(x)-S_n(x)]^2\mathrm{d}x。$$

因此,对 $f(x)$给出最佳近似的和 $S_n(x)$应使积分 J 达到最小值。从这个条件出发,贝塞尔确定了 $2n+1$ 个系数 a_0, a_1, …, a_n, b_1, …, b_n。因为我们将把 J 看做 $2n+1$ 个变元 a_0, …, b_n 的函数,取得最小值的必要条件是

$$\begin{cases}\dfrac{\partial J}{\partial a_0}=0,\ \dfrac{\partial J}{\partial a_1}=0,\ \cdots,\ \dfrac{\partial J}{\partial a_n}=0,\\ \dfrac{\partial J}{\partial b_1}=0,\ \cdots,\ \dfrac{\partial J}{\partial b_n}=0。\end{cases}\tag{11}$$

因为 J 是 a_0, …, b_n 的一个正二次函数,所以很容易看到,这 $2n+1$ 个方程所确定的变量的值,事实上将产生最小值。

如果在积分号内求微分，则方程组(11)成为

$$\begin{cases}\int_0^{2\pi}[f(x)-S_n(x)]\mathrm{d}x=0,\\ \int_0^{2\pi}[f(x)-S_n(x)]\cos x\,\mathrm{d}x=0,\ \cdots,\\ \int_0^{2\pi}[f(x)-S_n(x)]\cos nx\,\mathrm{d}x=0,\\ \int_0^{2\pi}[f(x)-S_n(x)]\sin x\,\mathrm{d}x=0,\ \cdots,\\ \int_0^{2\pi}[f(x)-S_n(x)]\sin nx\,\mathrm{d}x=0。\end{cases}\tag{11'}$$

现在，$S_n(x)$与正弦函数和余弦函数乘积的积分可以大大简化。对$\nu=0,1,\cdots,n$，我们有

$$\begin{aligned}\int_0^{2\pi}S_n(x)\cos\nu x\,\mathrm{d}x=\frac{a_0}{2}\int_0^{2\pi}\cos\nu x\,\mathrm{d}x+a_1\int_0^{2\pi}\cos x\cos\nu x\,\mathrm{d}x+\cdots\\ +a_n\int_0^{2\pi}\cos nx\cos\nu x\,\mathrm{d}x\\ +b_1\int_0^{2\pi}\sin x\cos\nu x\,\mathrm{d}x+\cdots\\ +b_n\int_0^{2\pi}\sin nx\cos\nu x\,\mathrm{d}x。\end{aligned}$$

根据角函数的初等积分性质，右端除去指标为ν的余弦项等于$a_\nu\cdot\pi$外，其余全为零，故有

$$\int_0^{2\pi}S_n(x)\cos\nu x\,\mathrm{d}x=a_\nu\cdot\pi\quad(\nu=0,1,\cdots,n)。$$

由于对a_0附有因子$\frac{1}{2}$，故此公式对$\nu=0$也成立。类似地，我们还有

$$\int_0^{2\pi}S_n(x)\sin\nu x\,\mathrm{d}x=b_\nu\cdot\pi\quad(\nu=1,\cdots,n)。$$

从这些简单的关系中得知，方程组(11′)中的每一个只含$2n+1$个未

知数中的一个，因此，能立即将它们的解写出来

$$a_\nu = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos \nu x\,\mathrm{d}x \quad (\nu = 0,\ 1,\ \cdots,\ n),$$
$$b_\nu = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin \nu x\,\mathrm{d}x \quad (\nu = 1,\ \cdots,\ n)。\tag{12}$$

如果今后取这些值作为 $S_n(x)$ 的系数，则 J 将达到最小值，且其值为

$$\int_0^{2\pi} f(x)^2\mathrm{d}x - \pi\left\{\frac{a_0^2}{2} + \sum_{\nu=1}^{n}(a_\nu^2 + b_\nu^2)\right\}。$$

值得指出，从我们最初对 $S_n(x)$ 所假设的形式得出的系数 a，b 和数 n 无关，而且在按同样原则逼近 $f(x)$ 时，项 $\cos \nu x$ 或 $\sin \nu x$ 的系数无论单独用这项或同时与任何其他项一起用，都具有完全相同的值。如果试图单独用余弦项 $a_\nu \cos \nu x$ 来获得 $f(x)$ 的最佳近似，即令

$$\int_0^{2\pi}[f(x) - a_\nu \cos \nu x]^2\mathrm{d}x = \text{最小值}。$$

我们会发现 a_ν 之值与我们在上面求得的一样。这个事实使得这种逼近方法特别便于使用。例如，如果由于一个函数相似于正弦函数而只用正弦级数来表示它，但发现其近似程度不够，还可以根据最小二乘法原则增加更多的项，而不需更改已找到的项。

我现在必须说明，用这样的方法确定的和 $S_n(x)$ 是如何实际趋向于 $f(x)$ 的问题，我们觉得最好模仿自然科学家所用的实验方法，先对若干具体情况画出几个近似曲线 $S_n(x)$，这就会使人对所发生的情况产生一个直观的印象，并要求作出数学解释，甚至使没有数学天赋的人也表现出兴趣来。

在以前的一次课程(1903—1904 年冬季学期)中，当我详细讨论这些问题时，我的助手席马克画过这种图，我会把其中某些原图放映

在屏幕上让你们看。

(1) 如果取由直线段构成的曲线,我们会得到所要求的简单而有启发的例子。例如,考虑在 $x=0$ 到 $x=\frac{\pi}{2}$ 与 $y=x$ 相同,在 $x=\frac{\pi}{2}$ 到 $x=\frac{3\pi}{2}$ 与 $y=\pi-x$ 相同,而在 $x=\frac{3\pi}{2}$ 到 $x=2\pi$ 一段又和 $y=x-2\pi$ 相同,然后周期地向 $(0,\ 2\pi)$ 之外扩展的曲线 $y=f(x)$。计算系数,会发现所有系数 a_ν 均为零,因为 $f(x)$ 是奇函数,所要求的级数形如

$$S(x)=\frac{4}{\pi}\left(\frac{\sin x}{1^2}-\frac{\sin 3x}{3^2}+\frac{\sin 5x}{5^2}-\cdots\right)。$$

在图 8.20 中画出了第一和第二个部分和的草图。部分和与曲线 $y=f(x)$ 的交点不断增加从而使之越来越近似。应特别注意的是,近似曲线越来越挤向曲线 $y=f(x)$ 在 $\frac{\pi}{2}$, $\frac{3\pi}{2}$, $\cdots$ 处的角的内部,虽然这些近似曲线作为解析函数本身不可能有角点。

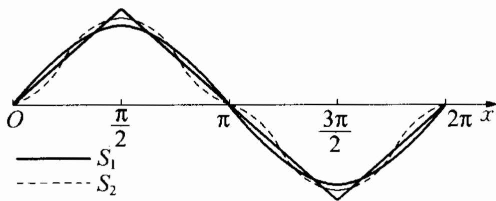


图 8.20

(2) 设 $f(x)$ 在 $x=0$ 到 $x=\pi$ 为 x,在 $x=\pi$ 到 $x=2\pi$ 为 $x-2\pi$,在 $x=\pi$ 处有一个跳跃。于是曲线由经过 x 轴上各点 $x=0$, 2π, 4π, $\cdots$ 的平行直线段组成。如果在间断点用垂直线段把不连续线段的端点连接起来,则函数将由一个不断开的折线来表示(图 8.21),看起来就像你们学写字时都写过的 m 的笔画。函数仍是奇的,所以余弦项消失,级数变为

$$S(x) = 2\left(\frac{\sin x}{1} - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \frac{\sin 4x}{4} + \cdots\right)。$$

图 8.21 画出了前二、三、四项之和。特别有趣的是注意它们是怎样模拟 $f(x)$ 的间断性的：在 $x=\pi$ 处越来越陡地通过零点。

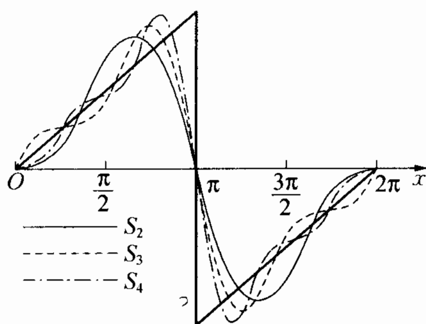


图 8.21

(3) 作为最后一个例子(图 8.22)，我将讨论一条曲线，它在 0 与 $\frac{\pi}{2}$ 之间等于 $\frac{\pi}{2}$，在 $\frac{\pi}{2}$ 和 $\frac{3\pi}{2}$ 之间等于 0，最后在 $\frac{3\pi}{2}$ 与 2π 之间等于 $-\frac{\pi}{2}$，然后周期地向外延展出去。如果在不连续处再次补上垂直线段，则得到一个钩状曲线。现在也只有正弦系数不为 0，因为函数是奇的，而级数为

$$S(x) = \sin x + 2\cdot\frac{\sin 2x}{2} + \frac{\sin 3x}{3} + 0 + \frac{\sin 5x}{5} + 2\cdot\frac{\sin 6x}{6}$$

$$+ \frac{\sin 7x}{7} + 0 + \frac{\sin 9x}{9} + \cdots。$$

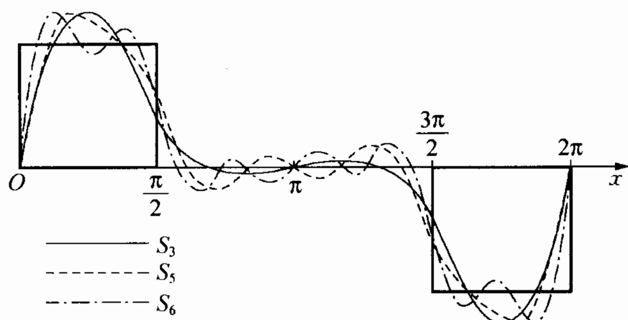


图 8.22

这里的系数规律不像以前那样简单,因此逐次的逼近曲线(图 8.22 画出了第三、第五、第六次近似)也不像以前的情况那样有那么类似的图像。

现在回到这样一个问题,即在一个确定的有限位置,若用 $S_n(x)$ 代替 $f(x)$,一般有多大误差? 到目前为止,我们只讨论了这个误差在整个区间上的积分。我们考虑系数 a_ν, b_ν 的积分公式(3),并用 ξ 表示积分变量以区别于用来表示一个确定点的 x。于是将有限和式(1)写成

$$S_n(x)=\frac{1}{\pi}\int_0^{2\pi}\mathrm{d}\xi\cdot f(\xi)\cdot\left[\frac{1}{2}+\cos x\cos\xi+\cos 2x\cos 2\xi+\cdots\right.$$
$$+\cos nx\cos n\xi+\sin x\sin\xi+\sin 2x\sin 2\xi+\cdots$$
$$\left.+\sin nx\sin n\xi\right],$$

或者将同一列加起来,可得

$$S_n(x)=\frac{1}{\pi}\int_0^{2\pi}\mathrm{d}\xi\cdot f(\xi)\left[\frac{1}{2}+\cos(x-\xi)+\cos 2(x-\xi)+\cdots\right.$$
$$\left.+\cos n(x-\xi)\right]。$$

方括号内的级数是很容易求和的,也许最方便的方法还是借用复指数函数。我在此不能详述,但如果利用被积函数的周期性,把积分区间改成从 $-\pi$ 到 $+\pi$,则得

$$S_n(x)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{d}\xi\cdot f(\xi)\,\frac{\sin\frac{2n+1}{2}(\xi-x)}{\sin\frac{1}{2}(\xi-x)},$$

为了判断这个积分的值,我们首先在 ξ 轴的区间 $x-\pi\leqslant\xi\leqslant x+\pi$ 上描出曲线

$$\xi = \pm \frac{1}{2\pi} \frac{1}{\sin \frac{1}{2}(\xi - x)}。$$

它们显然有类似于双曲线的分支(图 8.23)，且在这些分支之间，曲线

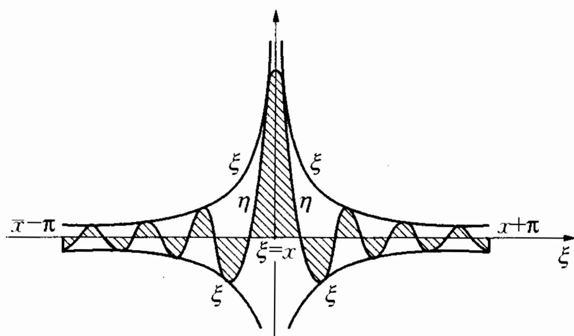


图 8.23

$$\eta = \frac{1}{2\pi} \frac{\sin \frac{2n+1}{2}(\xi - x)}{\sin \frac{1}{2}(\xi - x)} = \xi \cdot \sin \frac{2n+1}{2}(\xi - x)$$

来回振动，其频率随 n 变大而增加。当 $\xi = x$ 时，它的值是

$$\eta = \frac{2n+1}{2\pi}$$

并随 n 而增大。如果为了简单起见，令 $f(\xi) = 1$，则 $S_n(x) = \int_{-\pi}^{\pi} \eta \cdot d\xi$ 表示在曲线 η 和 ξ 轴之间的面积(图中阴影部分)。现在，任何一个对连续性稍有知识的人都会看到，如果 n 变得充分大，则右方和左方的面积都将正负相间，彼此将相互抵消，只剩下中间的狭长弧拱仍保留着。但很容易看到，随 n 的增加，它将理所当然地趋向于值 $f(x) = 1$。而且，只要 $f(x)$ 在 $x = \xi$ 处不振动得太强烈，一般情况下也是

如此。

无穷三角级数收敛的狄利克莱定理，其证明的基础就是这样一些考虑，而这些考虑是为了更正确使用而引申出来的。

狄利克莱于 1829 年在 *Crelle* 杂志第 4 卷上第一次发表了这个证明[①]。随后(1837 年)，他在达夫(Dove)和莫泽(Moser)编的 *Repertoriumder Physik* 中给出了一个更为人所熟知的形式[②]。现今，这个证明在大多数教科书上都有[③]，在此就不细述了。但我要讲使 $f(x)$ 能用无穷三角级数表示所必须满足的一些充分条件。再次设想 $f(x)$ 定义在 $0 \leqslant x \leqslant 2\pi$ 上，且周期地延拓出去。然后，狄利克莱提出了下面两个假设，现在简单地称为狄利克莱条件：

(a) 所给的函数 $f(x)$ 是逐段连续的，即它在区间 $(0, 2\pi)$ 上只有有限个间断点，在其他地方连续，直到跳跃点为止。

(b) 所给的函数 $f(x)$ 是逐段单调的，即可将区间 $(0, 2\pi)$ 分成有限个子区间，$f(x)$ 在每个子区间内或不增或不减。换句话说，$f(x)$ 只有有限个极大值和极小值(这就排除了像 $\sin\dfrac{1}{x}$ 那样在 $x=0$ 处极值有极限点的函数)。

在这些条件下狄利克莱证明了，在 $f(x)$ 的连续点，无穷级数精确地表示了函数 $f(x)$，即

$$\lim_{n\to\infty} S_n(x) = f(x),$$

而且狄利克莱还证明了，在间断点级数也收敛，但收敛于当 x 从两侧逼近间断点时 $f(x)$ 的极限值的平均值。这个结果通常表达为

① 1889 年柏林重印本，收于《狄利克莱论文集》第 1 卷，第 117 页。

② "Über die Darstellung ganz willkürlicher Funktionen durch Sinus-und Kosinusreihen"，重印后收于《狄利克莱论文集》第 1 卷，第 133—160 页，以及 1900 年莱比锡 *Ostwalds Klassiker* 第 116 期。

③ 参阅拜尔利著《傅里叶级数及球面调和》。

$$\lim_{n\to\infty} S_n(x) = \frac{f(x+0)+f(x-0)}{2}。$$

图 8.24 画出了这种间断点和对应的平均值。

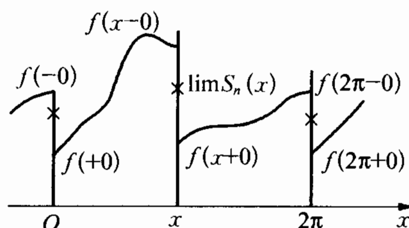


图 8.24

狄利克莱的这些条件对于 $f(x)$ 可以表示为级数 $S(x)$ 是充分的，但绝非必要条件。另一方面，只设 $f(x)$ 的连续性是不够的。事实上可以找到连续函数的例子，由于振动聚集得太厉害，级数 $S(x)$ 发散。

在介绍了这些理论方面的东西之后，我将回到三角级数的实用方面。要了解这里所产生问题的详细处理方法，我建议你们参考前面提到过的龙格写的书。你们会找到有关级数中系数的数值计算问题的详尽处理，即在给定函数的情况下，如何迅速地以最适当的方法计算 a_ν, b_ν 的积分。

为了计算这些系数，有人做了一种叫调和分析仪的特殊机械装置。叫这个名字，是由于展开函数 $f(x)$ 为三角级数与声学有关。这样的展开就是相应地把一个给定的音调 $y = f(x)$（x 表示时间，而 y 则为音调振动的振幅）分离为“纯音”，即纯余弦振动与纯正弦振动。在我们的收集品中，有苏黎世的卡拉第(Caradi)制造的分析仪，可以用它来确定 6 个余弦项和正弦项的系数 ($\nu = 1, 2, \cdots, 6$)，共 12 个系数。$\frac{a_0}{2}$ 必须用求积仪另外确定。迈克尔森(Michelson)等人曾做出一个能确定 160 个系数 ($\nu = 1, 2, \cdots, 80$) 的仪器，在朗格的书里讲了这种仪器。反之，这种装置也能把已给的含有 160 项的三角级数的和求出来，即从给出的系数 a_ν, b_ν 求函数。这个问题当然也有极大的实际意义。

迈克尔孙(Michelson)等人的仪器再一次使人们注意到一个十

分有趣的现象，这个现象早先曾有人注意过[①]，但很奇怪，过了几十年以后却被人忘却了，1899年吉布斯(J. W. Gibbs)在《自然》杂志上再次讨论了这个现象，从而被称为吉布斯现象。让我们对此现象说上几句。狄利克莱定理用公式$\frac{[f(x+0)+f(x-0)]}{2}$给出了无穷级数在$x$处的值。在上面讨论的第三个例子中(举例是为了心中记一个具体例子)，级数给出画在图8.25中函数在孤立点π, 3π, …处的值。

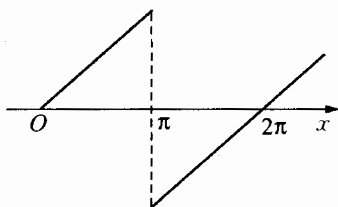


图 8.25

这样，我们用来解释三角逼近的方法就与狄利克莱的方法不同了。在狄利克莱的讨论中，x是固定的，而n趋于无穷。前面我们则设想n固定而x作为变量的$S_n(x)$，并逐次画出近似曲线$S_1(x)$, $S_2(x)$, $S_3(x)$, …。现在问：当n变成无穷大时，这些曲线发生了什么变化？或用算术的语言来说，当x作为变量而n变为无穷时，$S_n(x)$的极限是什么？从直觉来看，极限函数不会有上述那样的孤立点，而必然是一条连续的曲线。这条极限曲线似乎可能由$y=f(x)$的连续支加上在间断点处将$f(x+0)$和$f(x-0)$连接起来的垂直线段组成。在我们的例子中，极限曲线具有如图8.21所示的字母m的形状。但事实上，极限曲线的垂直部分投影却比$f(x+0)$和$f(x-0)$的范围超出一个有限量，因而极限曲线具有图8.26所画出的值得注意的形状。

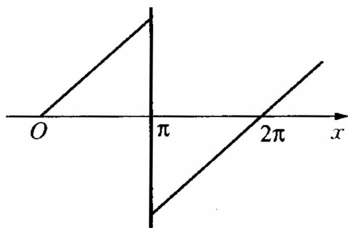


图 8.26

① 据《百科全书》第2卷第12条(Trigonometrische Reihen and Integrale)第1048页，H·威尔布里厄姆(H. Wibraham)已经熟悉了这里所讨论的现象，并从计算的观点进行了处理。

这个小小的重叠在曲线上方的塔是在迈克尔孙的仪器所描出的曲线上发现的，换句话说，它是实验所揭示的。起初认为这是机器不完善所致，但最后吉布斯认识到它是必然的。如果 $D=|f(x+0)-f(x-0)|$ 表示跳跃的高度，则根据吉布斯的计算，延伸长度为

$$-\frac{D}{\pi}\int_{\pi}^{\infty}\frac{\sin\xi}{\xi}\mathrm{d}\xi\approx\frac{1}{\pi}0.28D\approx 0.09D。$$

至于此结论的证明，只要对一个间断函数（例如我们的例子）给出就足够了，因为由它加上连续函数就可得到所有其他具有同样跳跃的函数。这个证明并不十分困难，可以立即从 $S_n(x)$ 的积分公式求得。而且，如果画出足够多的近似曲线，很清楚地可看到吉布斯点是怎样升起的。

如果我进一步考虑逼近曲线中许多有趣的微妙之处，那就说得太远了。我很高兴地建议你们参考费耶尔（Fejér）在《数学年刊》第64卷（1907年）中发表的一篇充实而非常容易读的文章。

我就此结束关于三角级数的专门讨论，并将漫步于另一个领域中，这个领域不论就其内容或历史都是与三角级数密切相关的。

4. 关于函数一般概念的插话

我很乐意在这些讲义中讨论函数概念，因为我们德国的中学教育改革运动都主张在教学中把这个重要概念放在一个突出的地位。

如果我们仍然遵循着历史的发展，我们首先注意到，像莱布尼兹和伯努利（Bernoulli）这些老一辈作家，都只对幂函数、三角函数这一类个别的例子使用函数概念。18世纪才出现一般的表述。

(1) 约于1750年，欧拉对函数这个词有两个不同的解释：

(a) 在他的“引论”中，他定义 x 的解析表达式如幂、对数、角函数等组成的每一个表达式为 x 的函数 y，但没有明确指出允许什么

样的组合。此外，他已经有了代数函数和超越函数这种人们熟知的区分。

(b) 与此同时，他定义函数$y(x)$为 xy 坐标系上随手画出的(libero manus dnctu)一条曲线(图 8.27)。

(2) 在 1800 年左右，拉格朗日在《解析函数论》中把函数概念限于所谓的解析函数。即能用 x 的幂级数定义的函数，这与欧拉的第二个定义恰成对照。在现代用语中，解析函数保留了同样的含义。当然必须承认，这只包含分析中实际出现的函数的一个特殊类别。一个幂级数

$$y = P(x) = a_0 + a_1x + a_2x^2 + \cdots$$

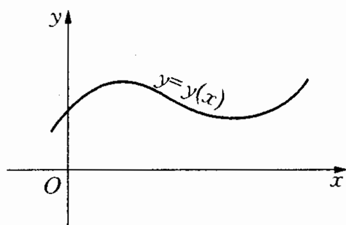


图 8.27

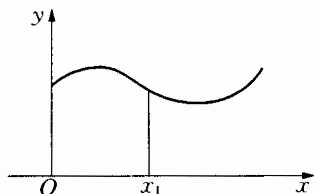


图 8.28

只在它的收敛域即某个围绕 $x = 0$ 的区域内定义了一个函数。但是，不久后发现了一个方法，将函数的定义范围扩大到这个区域以外。如果 x_1 是在 $P(x)$的收敛区域内(图 8.28)，且 $P(x)$被分解为$(x - x_1)$ 的幂的新级数

$$y = P_1(x - x_1),$$

它有可能在一个超过前一个范围的区域里收敛，所以在一个较大的区域内定义了 y。重复这个步骤，可能进一步扩大这个区域。熟悉复变函数理论的人都知道这个解析延拓法。

特别要注意的是，只要知道函数 y 在 x 轴的任意一小段，例如在 $x = 0$ 的某个邻域内的性质，幂级数 $P(x)$的每个系数，从而整个

函数 y 就被确定了。因为此时 y 在 $x=0$ 的各阶导数值就知道了，而我们知道

$$y(0)=a_0,\ y'(0)=a_1,\ y''(0)=2a_2,\ \cdots。$$

因此，一个拉格朗日意义上的解析函数的全部形态，可由它在一个任意小段内的形状完全确定。这个性质和欧拉第二个定义下的函数的性质是完全相反的。按欧拉的定义，曲线的任何部分，可以按人们的意愿任意地连续延伸。

(3) 函数概念的进一步发展要归功于傅里叶，他是 19 世纪初在巴黎工作的众多的重要数学家之一。他的主要著作是 1822 年出版的 *Théorie analytique de le chaleur*（《热的解析理论》）。但他在 1807 年就第一次把他的理论通告了巴黎科学院。这部著作是当今数学物理中用得较多的一个影响深远的方法起源，这个方法可以刻画如下：将一切问题归结为附有初始条件的偏微分方程的求解，归结为所谓的边值问题。

傅里叶特别处理了热传导问题，其中一个简单情况，可叙述如下。一块圆板的边界被保持在恒温下，例如一部分处于冰点；而另一部分处于沸点（图 8.29）。由此产生的热的流动，最终将给出什么样的定常温度呢？这里在边界的各部分可互不相关地给出边界值。因此，欧拉的第二个函数定义，比之拉格朗日的定义而言，更适合于这里的情况。

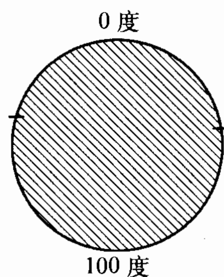


图 8.29

在前面提及的狄利克莱的著作中，基本上保持了这个定义，只是改用了分析语言，或用现代术语来说是算术化了，这样做事实上是必要的。因为，一条曲线无论画得怎样精细，都不能准确确定 x 和 y 的对应关系。钢笔的笔画总有一定宽度，因而相互对应的 x 和 y 的长度只能测到一定小数位的精确程度。

(4) 狄利克莱用下面方法将欧拉的定义算术化：如果对应于给

定区间内的每一个 x 值，有一个确定的 y 值被确定，则 y 称为 x 的一个函数。虽然他发表了函数的这个最一般的概念，但他经常想到的主要还是连续函数，或不太间断的函数，因为那时一般都是这样做的。人们认为，间断点聚集在一起的复杂情形是可以想象的，但他们未必认为值得如此注意。狄利克莱甚至在十分精确地叙述他所考虑的一切方程皆须满足的狄利克莱条件时，也只说把"完全任意的函数"（和傅里叶所说的"fonctions entièrement arbitrairies"一样）展开为级数，就表现了这种立场。

（5）现在，我们必须讲到这样的事实，大约在 1830 年左右，复变函数理论开始独立发展了，且在随后的 30 年里，变成了数学家的共同财富。这个发展首先是与柯西、黎曼和魏尔斯特拉斯的名字联系在一起的。你们知道，前两个人是从复变函数

$$f(x+\mathrm{i}y)=u+\mathrm{i}v$$

的实部与虚部 u、v 必须满足的冠以他们名字的偏微分方程出发；而魏尔斯特拉斯则用幂级数和他的解析延拓的总体来定义函数，因而，在一定意义上来说，他是追随拉格朗日的。

很值得注意的是，进入复数领域之后，上面考虑过的两种函数概念一致且联系了起来。我来简单讲一下。

令 $z=x+\mathrm{i}y$，并考虑幂级数

$$f(z)=u+\mathrm{i}v=c_0+c_1z+c_2z^2+\cdots。\qquad(13)$$

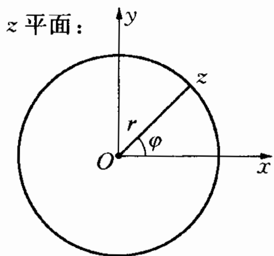


图 8.30

当 $|z|$ 很小时它是收敛的，用魏尔斯特拉斯的话来说，它确定了一个解析函数的元素。考虑它在以 $z=0$ 为中心的，整个位于收敛域内的、充分小半径 r 的圆内的值，即在幂级数里，令 $z=x+\mathrm{i}y=r(\cos\varphi+\mathrm{i}\sin\varphi)$（图 8.30），则有

$$f(z)=c_0+c_1r(\cos\varphi+\mathrm{i}\sin\varphi)$$

$$+c_2r^2(\cos 2\varphi+\mathrm{i}\sin 2\varphi)+\cdots。$$

如果将系数分成实部与虚部

$$c_0=\frac{\alpha_0-\mathrm{i}\beta_0}{2},\ c_1=\alpha_1-\mathrm{i}\beta_1,\ c_2=\alpha_2-\mathrm{i}\beta_2,\ \cdots,$$

则得到 $f(z)$的实部为

$$u=u(\varphi)=\frac{\alpha_0}{2}+\alpha_1 r\cos\varphi+\alpha_2 r^2\cos 2\varphi+\cdots$$
$$+\beta_1 r\sin\varphi+\beta_2 r^2\sin 2\varphi+\cdots。\tag{14}$$

为了使(14)式中所有符号为正,把 c 的虚部符号取成负的。因此,$f(z)$的幂级数在圆周上给出实部 u 的值,此值是角 φ 的函数,而幂级数则完全和以前的三角级数一样,它的系数是 α_0, $r^\nu\alpha_\nu$, $r^\nu\beta_\nu$。

当然,只要圆(r)整个地在幂级数(13)的收敛域内,这些值 u 就是拉格朗日意义上的解析函数。但是,如果这个圆就是级数(13)的收敛域的边界圆,则级数(13)因而还有级数(14),就不一定收敛了。也可能级数(14)继续收敛,此时边界值 $u(\varphi)$可能不是狄利克莱意义上的解析函数。

如果我们反过来做,对圆(r)指定一个满足狄利克莱条件的值 $u(\varphi)$的任意分布,那么,它们可以展开为形式(14)的三角级数,从而量 α_0, α_1, $\cdots$, β_1, β_2, β_3, $\cdots$及幂级数(1)的各系数$\left(除了相差一个任意常数\dfrac{-\mathrm{i}\beta_0}{2}外\right)$都被确定。可以证明,此幂级数在圆$(r)$内真正收敛,而它确定的解析函数的实部在圆$(r)$上以 $u(\varphi)$为边值。或者说得更准确些,当自变量趋近于 $u(\varphi)$的连续点 φ 时,这个实部趋向于值$u(\varphi)$。

这些证明,全被纳入关于收敛圆上幂级数性质的研究中。当然,我在此不再详述。但从这些说明中也许能看出,傅里叶—狄利克莱函数概念与拉格朗日的函数概念是以何种方式彼此融合在一起的,

即圆的边界上的三角级数 $u(\varphi)$ 的性态有任意性，而对幂级数，这种任意性只集中在紧接圆心的邻域内。

(6) 现代科学没有停止在这些概念的表述上面，即使个别研究者感到厌倦了，科学也是永不停息的。在最近的 30 年中，数学家采取了与狄利克莱非常不同的观点，抓住具有尽可能最大的间断性的函数，特别是不满足狄利克莱条件的函数，已经找到了最引人注目的若干类函数，它们具有最令人讨厌的“团成可怕的一堆”的奇异性。于是便产生了这样一个问题，即对“合乎常理”的函数成立的一些定理，对这种变态的函数又有多大可靠性。

(7) 与这一问题相联系，最后又提出了函数概念更深远的推广。到目前为止，一般都认为，函数是定义在由所有的实值或复值的 x 组成的连续统上的，或至少在一个完整的区间或区域内的每一点都有定义。但最近由康托发明的点集理论越来越受到注意。据此理论，所有的 x 所组成的连续统只不过是点集的一个明显的例子。从这个新观点出发，只在某个任意集合的元素 x 处有定义的函数也要考虑。所以，若对某种东西(数或点)的集合 x 的每个元素，都有集合 y 的一个元素相对应，则称 y 为 x 的函数。

让我指出这个最新的发展与原来的概念之间的差别。从第(1)点到第(5)点所考虑的概念，主要都是结合自然界中的应用而提出和发展起来的。只要想一下傅里叶著作的题目就知道了。但第(6)点和第(7)点中所述的新研究，纯粹是从对数学研究的爱好出发而得到的结果，它并不考虑自然现象的任何需要，其结果也确实尚未找到直接应用。当然，乐观者会认为，有朝一日总会得到应用。

现在回到我们习惯的问题：以上种种概念，有多少应放到中学里来学？教师和学生又应该掌握多少？

在这方面我想说，学校教育总是落后于最新的科学进展相当的一段时间，起码几十年，这种情况不仅是情有可原，甚至也是符合需要的。就是说，应该有某种差距。但现在实际存在的差距，在某些方

面不幸太大了。它超过了一个多世纪，因为在绝大多数学校里都忽视了欧拉时代以来的整个发展。因此，有很大的改革余地。把我们的改革要求和科学的现状相比，实在也所望不奢。我们仅希望，根据欧拉对函数的某一个解释，把函数的一般概念作为一种发酵物掺和到整个中学的数学教学中去。当然，不应该用抽象的定义来介绍它，而应通过初等的例子，向学生作主动的介绍，这种例子在欧拉那里是很多的。但对于数学教师，要求应更高一些，至少要知道复变函数论的初步知识；虽然我不同样要求教师掌握点集理论的最新概念，但仍然希望在众多的教师之中有一小部分人独立钻研这些问题。

我愿就三角级数理论在整个发展中曾经起过的重要作用再说上几句。你们可以在伯克哈特写的 *Entwickelungen nach oszillierenden Funktionen*（特别是第 2、第 3、第 7 章）中找到大量的有关文献。正如伯克哈特的朋友所说，这是一本“大部头的报告”，自 1901 年后，在 *Jahresbericht der deutschen Mathematikervereinigung* 第 10 卷中连载[①]，其中有 9 000 多条文献，你们在任何地方都难以找到这么多编在一起的有关文献。

第一个用三角级数来表示一般函数的是约翰·伯努利(Bernoulli, J.)的儿子丹尼尔·伯努利(Bernoulli, D.)。1750 年左右，他在研究弦振动声学问题时，指出弦的一般振动可以用对应于基波和谐波的正弦振动的叠加来表示。这正涉及把表示弦振动的函数展开成它的三角级数。

虽然不久后级数知识有了进步，但仍然没有一个人真正相信，用图形给出的任意函数可以用级数来表示。归根到底是因为存在着一种不确定的预感，觉得还要再作考虑。现在，通过点集理论，我们已

① 该报告以两卷半的篇幅完成，作为第 10 卷第二分册出版，莱比锡，1908 年。后在《数学科学百科全书》第 2 卷中发表了一个简要的摘录。又：伯克哈特的报告收录到 1850 年为止，1850 年以后的发展，由希尔伯(Hilb)和里茨(Riesz)撰文在《数学科学百科全书》第 2 卷(C10)中作了简单介绍。

十分清楚是怎么回事了。也许有人曾经假设(自然,不见得能将感觉精确表述出来),一切任意函数的集合,即使将不连续函数除外,也大于表示全部三角级数的所有可能数组 $a_0, a_1, a_2, \cdots, b_1, b_2, \cdots$ 的集合。

只有现代点集理论的精确概念才能澄清这团疑云,并证明这个论断是错误的。让我借此机会对这个重要之点作一番说明。很容易看到,一个定义在给定区间,例如 0 到 2π 的任意连续函数,如果在所述区间的所有有理点上的值已给定,则该函数的整个形状就确定了下来(图 8.31)。因为这些有理点是稠密的,所以可以通过函数在有理点的值来任意逼近在无理点的值,由于函数的连续性,函数 $f(x)$ 的值可以用函数在逼近点处的值求极限来确定。进而,我们知道,所有有理点之集合是可数的(见附录Ⅱ),即能把它们排成无穷数列,即有一个确定的第一元素,后边接着是确定的第二个元素,再后边是确定的第三个元素,如此等等。由此可知,确定任意连续函数,无非是确定一个适当的可数的常数集,即函数在可编序的有理点上的值。用同样的方法,即通过常数的可数序列 $a_0, a_1, b_1, a_2, b_2, \cdots$,我们可以确定一个三角级数。因此,怀疑连续函数的总体按其本性是否真正大于级数的总体是没有根据的。对满足狄利克莱条件的不连续函数,也可作类似考虑。后面我们还有机会详谈。

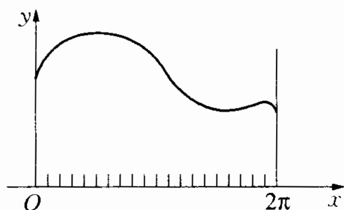


图 8.31

把这些疑虑一下子统统扫到一边的人是傅里叶,而也正因为如此,使他在三角级数历史上成了如此重要的人物。当然,他没有把结论建立在点集理论的基础上,但他是第一个有勇气相信可以用级数来表示函数的人。在这个信念的鼓舞下,他像我们在前面所做的那样,用若干不连续函数的典型例子,通过实际计算建立了许多级数。收敛性的证明是后来由狄利克莱给出的,他也是傅里叶的学生。傅

里叶的这个立场有革命性的影响。对于在不同区间遵从完全不同规律的任意函数，有可能用解析函数的级数来表示，对于当时的数学家来说，这是完全新的东西，也是他们所意料不到的。为了承认傅里叶揭示了这种可能性，所以在他所研究的级数之前冠上了他的名字，并一直持续到今天。

最后我要简单地提一下傅里叶的第二项功劳。这就是，当被表示的函数的周期可以变成无穷时，他考虑了三角级数的极限情形。因为一个具有无限长周期的函数就是一个任意的定义于整个 x 轴的非周期函数，所以这种极限的情形为表示非周期函数提供了一个方法。对级数的自变量作线性变换，即可用 l 代替 2π 作为函数的周期，然后令 l 趋向无穷。级数于是变成所谓傅里叶积分

$$f(x)=\int_0^{\infty}[\varphi(\nu)\cos\nu x+\psi(\nu)\sin\nu x]\mathrm{d}\nu,$$

其中 $\varphi(\nu)$，$\psi(\nu)$ 以一定的方式表示为 $f(x)$ 从 $-\infty$ 到 $+\infty$ 的积分。这里的新东西是，指数 ν 连续地取从 0 到 ∞ 的所有值，而不是只取 0，1，2，…；同时，$\varphi(\nu)\mathrm{d}\nu$，$\psi(\nu)\mathrm{d}\nu$ 相应地取代了 a_ν 和 b_ν。

迄今为止，我们在讲分析时主要涉及初等超越函数，现在我们把它放下，进入新的、最后的一章。

第九章
关于无穷小演算本身

当然我假设你们都知道如何微分与积分，且常常使用这两种方法。这里只涉及较为一般的问题，例如逻辑与心理学基础问题、教学问题等等。

§9.1 无穷小演算中的一般考虑

我先对数学的范围讲一些一般的初步看法。你们常常会听到非数学家，特别是哲学家说，数学仅仅是由清楚叙述的前提得出的结论，在这个过程中，并不过问这些前提意味着什么，是真还是假，只要不相互矛盾就行了。但在数学方面做过创造工作的人说法完全不同。事实上，前面那种人只想到已经完成的数学理论所铸成的结晶形式。但是，数学研究工作者也和其他科学研究者一样，往往并不按照严格演绎的方式行事。相反，他基本上用的是想象的办法，借助于试探的手段，不断归纳前进。有的数学家发现了十分重要的定理，但却未能给出证明，这样的例子是很多的。然而，能否认这是伟大的成就吗？能苟同前面的定义，硬说这不是数学，只有提供了完整证明的后继者才做了真正的数学工作吗？无论怎样说，如何使用数学这个词是任意的事。但没有一种价值标准能否认第一个宣布定理的人所做的归纳工作，至少和第一个证明定理的人所做的推导工作一样有价值。因缺一不可，发现是随后作结论的先决条件。

在微积分的发现和发展过程中，归纳过程就起了这样一种重要作用，其中并无非此不可的逻辑步骤，而有效的幻想手段，又常常是

意识的直觉作用。我这里讲的是直接的直觉，尽管有各种不确切之处。例如讲到一条曲线，我立即想到的是有一定宽度的笔画，而不是一个抽象的概念，不需要假设极限过程才产生一维曲线。我想通过微积分思想的历史发展情况来证实这个观点。

首先拿积分概念来说，历史上它起源于测量面积与体积（求面积与体积）问题。抽象的逻辑定义把积分 $\int_a^b f(x)\mathrm{d}x$ 定义为介于曲线 $y=f(x)$，x 轴，和直线 $x=a$，$x=b$ 之间的面积，即内接于这个区域的狭长矩形面积之和在数目无限增加、矩形宽度无限减小时的极限。然而，感官的直觉却自然地把这个面积定义为大量狭长矩形面积之和，而不是这个严格的极限。事实上，画图必然有不准确处，不可避免地使矩形不能画得太窄（图 9.1）。

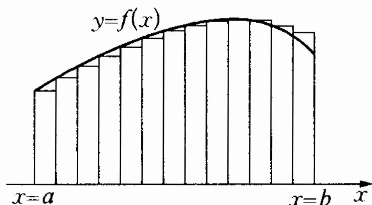
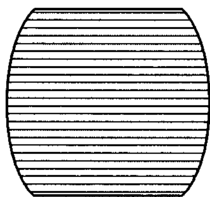


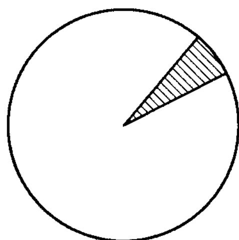
图 9.1

这个朴素的想法，就是无穷小演算早期最伟大的研究者的想法。让我首先谈谈开普勒（Kepler）。他在“Nova stereometria doliorum vinariorum”[①]中谈到了物体的体积。他的主要兴趣是测量木桶，以确定木桶的最合适形状。他完全采用了上述朴素的观点。他把木桶的体积看成由一层层适当排列的许多薄片组成（图 9.2），并把它看做是这些薄片的体积之和，每一片是一个圆柱体。他用同样的方法计算了简单的几何体，例如球的体积。他设想球由许多在球心处有共同顶点的小棱锥体组成（图 9.3）。根据众所周知的棱锥体积公式，它的体积将是所有棱锥的底面积之和的 $\frac{r}{3}$ 倍。把这些小面积之

---

① “Linz on the Danube”（“多瑙河上的林茨”），1615 年。德文发表于 *Ostwalds Klassiker*，第 165 期，莱比锡。


图 9.2


图 9.3

和简单地写成球的表面积，即 $4\pi r^2$，他得到体积的正确公式 $\frac{4\pi r^3}{3}$。不但如此，开普勒还明确地强调了这种想法的实际启发价值。至于严格的数学证明，他引用所谓穷举法。例如，这个曾被阿基米德用过的方法，是这样求圆的面积的：借助圆内接与圆外切的正多边形，不断增加边数，小心地求出多边形面积，从而确定圆面积。它和现代方法之间的本质差异在于它默认不言而喻地存在一个数，以度量圆面积。而现代的无穷小演算却不肯接受这个直觉的证据，而使用抽象的极限概念，把这个数定义为度量内接正多边形面积的许多数的极限。然而，尽管这个数的存在是假定的，但穷举法通过已知的直线图形面积去逼近未知面积的过程是严密的，符合现代的严格要求。不过这个方法往往十分麻烦，不适合于发现面积和体积。海伯格(Heiberg)在 1906 年发现的一篇阿基米德的文章表明，阿基米德事实上在研究中并没有用穷举法。在通过其他方法获得结果后，他之所以用穷举法作出证明，是为了适应那个时代对严格性的要求。他发现他的定理时所用的方法，考虑到了重心、杠杆定律，也包括直觉在内，例如，用平行弦序列组成三角形和抛物线弧形，用平行圆盘序列组成圆柱、球和圆锥。

现在回到 17 世纪，我们在耶稣会教士卡瓦列里(Bonaventina Cavalieri)的 *Geometria indivisibilibus continuorum nova quadam*

ratione promota[①]书中找到了类似于开普勒的思想。在那本书里，卡瓦列里建立了今天以他的名字命名的原理："两个物体，如果它们到底面的相等距离处的截面面积相等，则体积相等。"[②]你们都知道，这个卡瓦列里原理是中学里常用的一个原理。人们曾相信用这种办法可以回避积分。但事实上，这个原理本身就完全属于微积分学。卡瓦列里建立这个原理时设想，两个立体都是由薄片层叠加起来的，这些薄片是成对相同的，所以能够通过移动一个立体的每一片而把它变成另一个立体(图 9.4)。这样做不会改变体积，因为移动前后的薄片都是一样的。

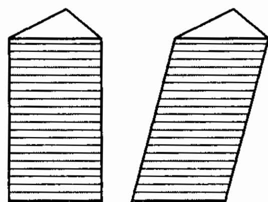


图 9.4

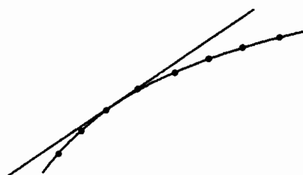


图 9.5

函数的微商，即曲线的切线，同样是凭朴素的感官直觉导出的。在这种情况下，是用一条折线来代替曲线(实际上就是这样做的，见图 9.5)，该折线取曲线上足够多的相互靠近的点作为顶点。由于我们的感官的直觉的本性，很难把曲线同这些点的集合区分开来，更不用说同折线区分开来了。于是切线干脆定义为连接两个相邻点的直线，即折线的一个边的延线。从抽象的逻辑观点来看，无论点是如何靠近，这条线仍然只是割线；只有当两点距离趋向于零时，割线趋向的极限位置才是切线。此外从这种朴素的观点来看，曲线的曲率圆又被看做通过折线的顺序三顶点的圆，而严格的过程则定义曲率圆为所述三点无限接近时该圆的极限位置。

---

① 博诺尼阿，1635 年。1653 年第一版。

② 我国称为祖暅原理。——中译者

对这种朴素的直觉信服的程度，当然因人而异，很多人，包括我自己在内，对此感到满足。另一些只有纯逻辑才能的人，则认为直觉是毫无意义的，而且无法想象会有人把直觉当作数学思想的基础，然而这类考虑常常成为新的、富有成效的探索的起点。

而且这种朴素的方法，常常在一个数学物理、力学或微分几何的预备定理被建立时，起着不知不觉的重要作用。你们都知道，它们在那种情况下是十分有用的。当然，纯粹的数学家对这些情况是不屑一顾的。在我当学生的那个时候，有人说，物理学家把微分当成一个铜片，像对待其他仪器一样。

这一方面，我愿称赞一下莱布尼兹的记号。这个记号之所以在今天通用，是因为它既适当考虑了朴素的直观，又一定程度地照顾到隐含在概念里的抽象极限过程。例如，莱布尼兹的微商记号$\frac{\mathrm{d}y}{\mathrm{d}x}$，首先提醒人们回忆起它来源于商。但与有限差的常用符号Δ相对比，d表示增加了某些新的内容，即过渡到极限。同样，积分记号$\int y\mathrm{d}x$首先暗示积分来源于微小量的和，但没有用表示和的常用记号$\sum$，而用拉长了的S[①]，以表示在求和过程中还加了些新内容。

现在稍微深入地讨论一下微积分学的逻辑基础，同时讲讲微积分历史发展情况。

(1) 大学里通常所教的主要思想在于说明微积分只是一般极限概念的应用，微商定义为自变量与函数的相应有限增量的商的极限(倘若此极限存在)

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x},$$

而不把微商看做是具有独立意义的量$\mathrm{d}y$和$\mathrm{d}x$的商。同样地，积分

---

① 值得注意的是，许多人不知道$\int$有这种意思。

定义为和的极限

$$\int_a^b y\mathrm{d}x = \lim_{\Delta x_i \to 0} \sum_{(i)} y_i \cdot \Delta x_i,$$

其中 Δx_i 是区间 $a \leqslant x \leqslant b$ 的有限子区间，y_i 是函数在该子区间内的任意值，且所有 Δx_i 趋向于 0。但 $y\mathrm{d}x$ 并没有诸如和的被加数一类的任何实际意义。这些记号保留至今，是由于前面讲的方便的缘故。

(2) 如上所述的概念，是由牛顿自己准确地建立起来的。请看他在 1687 年发表的基本著作《自然哲学的数学原理》[①]中的一段原话："Ultimae rationes illae, quibucscum quantitates evanescunt, revera non sunt rationes qnanitatum ultimarum, sed limites, ad guos quantitatum sine limite descrescentium rationes semper appropinguant, et quos propius assequi possunt, guam pro date guavis differentia, nunguam vero transqredi negue prius attingere quam quantitates diminuuntar in infinitum."此外，牛顿回避无穷小演算，尽管实际上在推导结果时用到了它。因为他提出无穷小演算方法的基本著作写于 1671 年，直到 1736 年才发表，那本书名为 *Methodus Flexionum et Serierum Infinitarum*[②]。

在后一本书里，牛顿用很多例子阐明了这种新方法，但没有对方法的基础作解释。他联系到一个日常生活现象，这个现象暗示着求极限，即：若考虑在 x 轴上时刻 t 的位移为 $x = f(t)$ 的运动，则每个人对运动速度指的是什么都已有一个概念。如果我们分析这个运动，就会发现原来是指差商 $\frac{\Delta x}{\Delta t}$ 的极限值。牛顿把这个 x 对于时间 t 的速度作

---

① W·汤姆逊(W. Thomson)及 H·布莱克本(H. Blackburn)编的新版本，格拉斯哥，1871 年，第 38 页。

② I·牛顿，*Opuscula Mathematica, philosophica, et philologica*，第 1 卷，第 29 页，洛桑，1744 年。

为他研究的基础。他称为 x 的流数，并记为$\dot{x}$。他把变量 x，y 都看做是依赖于时间 t 这个基本变量的，于是微商$\frac{\mathrm{d}y}{\mathrm{d}x}$作为两个流数的商$\frac{\dot{y}}{\dot{x}}$出现，现在我们则应把它记为较完全的形式$\left(\frac{\mathrm{d}y}{\mathrm{d}t}:\frac{\mathrm{d}x}{\mathrm{d}t}\right)$。

(3) 牛顿的这些思想被 18 世纪的许多数学家所继承和发展，他们在极限概念的基础上，以不同程度的严格性建立了无穷小演算，这些人中我只举出几个，麦克劳林(C. Maclaurin)、达朗贝尔以及哥廷根大学的卡斯特纳等。大家可以参阅麦克劳林的《流数论》[①]，这本书作为一本教科书，当然有很广的影响。还可参阅达朗贝尔的伟大的《方法论百科全书》，以及卡斯特纳的讲义和书[②]。欧拉基本上也属于这些人之列，尽管在他身上还表现出其他一些倾向。

(4) 为使无穷小演算可以称得上是一个前后一致的体系，必须填补以上这些发展中的一个关键性的空白。确实，微商已被定义为极限，但还缺少用它来估计函数在一个有限区间内的增量的方法。中值定理提供了这个方法。柯西的伟大贡献在于承认它的根本重要性和使它随之成为微分学的出发点，所以称柯西为现代意义上严格的无穷小分析的奠基人是不过分的。他在这一方面的基本著作是根据在巴黎的讲稿写成的 *Résumé des Leçons sur le Calcul Infinitésimal*[③]，该书第二版只公布了第一部分 *leçons sur le Calcul Differentiel*[④]。

如你们所知，中值定理可以表述如下：如果一个连续函数 $f(x)$在给定区间内有微商 $f'(x)$，则在 x 与 $x+h$ 之间必有一点 $x+\theta h$，使得

$$f(x+h)=f(x)+h\cdot f'(x+\theta h)\quad(0<\theta<1)。$$

---

① 爱丁堡，1742 年。

② A · G · 卡斯特纳：*Anfangsgründe der Analysis des Unendlichen*，哥廷根，1760 年。

③ 巴黎，1823 年。《全集》第 4 卷第 2 分册，巴黎，1899 年。

④ 同上。

注意这里出现的 θ 是中值定理所专有，也是使初学者开始常常感到迷惑不解的。这个定理的几何意义是十分明显的。它仅表示，在曲线上的点 x 与 $x+h$ 之间，有一点 $x+\theta h$，其切线平行于连接点 x 和 $x+h$ 的割线(图 9.6)。

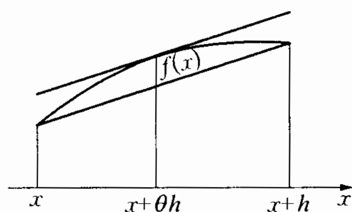


图 9.6

(5) 如何对中值定理给出一个严格的算术化证明而不求助于几何直观呢？当然，这种证明只能意味着把定理推回到变量、函数、连续等的算术化定义，因而也就必须把这些定义预先以抽象和准确的形式建立起来。这种严格的证明只好等到魏尔斯特拉斯和他的追随者来做了。数的连续统的现代算术化概念的传播也要归功于这些人。我将向你们介绍他们论据的关键之点。

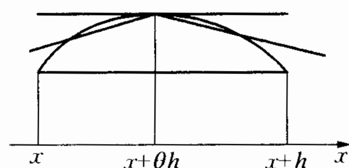


图 9.7

首先，很容易把这个定理化成割线是水平的情况，即 $f(x)=f(x+h)$ (图 9.7)。于是，必须证明存在一点，其切线是水平的。为此，可以利用魏尔斯特拉斯定理：每一个在闭区间内连续的函数，在区间里至少取得一次极大值和极小值。据我们的假设，所给函数除非是个常量，至少有一个极值在区间$(x, x+h)$内部取得。让我们设在$x+\theta h$处取得极大值(极小值情形可用同样方法处理)。于是 $f(x)$ 在$x+\theta h$ 的邻域内不可能取得比此点的值更大的值，所以右差商是负数或是零，左差商是正数或是零。因按假设，在区间内每一点的微商都存在，它在 $x+\theta h$ 的值可以被当作为一些值的极限，这些值或者是非正的或者是非负的。因此，它的值必定是零，在 $x+\theta h$ 处的切线是水平的。定理证毕。

今天科学的数学是按我们所概述的发展顺序建立起来的。但许

多世纪以来，还有一个完全不同的无穷小演算概念平行地发展着。

(1) 上述情况可回溯到古代关于连续统的构造的形而上学的思想，认为连续统是由最终不可分的无穷小部分构成的。在古代就有了这种不可分量的暗示，后来则由经院派学者加以广泛的研究，特别是耶稣会教士们更把它推进了。作为一个典型例子，我们回忆卡瓦列里的 *Geometrin Indivisibilibus Continuorum Promota*，这本书指明了它的真正本性。事实上，他只不过是用间接的方法考虑了直观的数学近似。他把空间看做是由最终不可分的部分"indivisibilin"所组成。在这方面，了解数百年间乃至上千年间对连续统概念的种种分析是有意义和重要的。

(2) 与牛顿分享发明无穷小演算的盛名的莱布尼兹，也采用了这种思想。对他来说，最基本的东西并不是看做一个极限的微商。变量 x 的微分 dx 对于他是真正存在的，是横轴的最终的不可分部分，它是异于零但小于任何有限量的一个量(即所谓"真正的"无穷小)。同样地，高阶微分 d^2x, d^3x, …定义为二阶无穷小量、三阶无穷小量，…它们的每一个"与前一个相比都是无穷小量"。因此，有了一串不同级别量的系统。据不可分量的理论，由曲线 $y=f(x)$ 和横坐标轴所包围的面积，是所有个别的纵坐标之和。正是基于这一观点，莱布尼兹在积分学的第一本手稿(1675 年)里，写的是 $\int y$ 而不是 $\int y dx$。

但是，莱布尼兹感兴趣的绝非仅仅是这个观点。有时他使用数学的近似概念，例如说微分 dx 是一个有限线段，但小到这样的程度，使曲线在此区间内与切线没有值得注意的差别。上述形而上学的思想，肯定只是这些简单的心理因素的理想化。

但是莱布尼兹的数学思想还有第三个方面，而且是他所独有的，这就是他的形式化观点。我常常提醒你们，可以把莱布尼兹看做形式数学的奠基人。他的思想如下所述：不管我们对微分赋予什么意义，甚至不管是否加上任何意义，只要对它们确定适当的运算规则，

并适当地加以应用，必然会得到某些合理与正确的结果。莱布尼兹反复地提到了微分与复数的相似性，他对复数也有相似的想法。至于这些微分运算的规则，他主要用到公式

$$f(x+\mathrm{d}x)-f(x)=f'(x)\mathrm{d}x。$$

中值定理表明，只有用 $f'(x+\theta\mathrm{d}x)$ 代替 $f'(x)$时才是正确的，但写成 $f'(x)$所造成的误差是高阶（二阶）无穷小，而在对微分作计算时（这是最重要的形式规则）可以忽略。

莱布尼兹发表的主要作品包含在著名的第一本科学杂志 *Acta Eruditorum*[①] 中，分别发表于 1684 年、1685 年、1712 年。你们会发现，在第一卷中以"Nova methodus pro maximis et mimimis"（第 467 页及以后部分）为标题第一次公布了有关微分的研究。在这篇论文中，莱布尼兹仅仅发展了微分法的规则。以后的论文还阐明了原理，主要着重于形式的观点。在这一方面，他的即将逝世前的 1712 年发表的那篇短文，是特别具有典型意义的[②]。其中，他直率地谈到了那些仅仅是普普通通的（toleranter vera 或法文 passables）定理及定义："Rigorem quidem non sustinent, habent tamen usum magnum in calculando et ad artem lnveniendi universalesque conceptus valent."他这里所指的，既指无穷小，又指复数。如果我们也许谈到无穷小，那么，"commoditati expressionis seu breviloquio mentalis inservimus, sed non nisi toleranter vera oquimur, quae explicatione rigidantur"。

（3）自莱布尼兹以后，这种新的演算以他的思想为中心，在欧洲大陆迅速得到了普及，而且他的 3 个观点都可以找到各自的代表。这里，我必须提到以前所不曾出现过的第一本微分教科书：马奎

① 有一部分已译成德文，发表于 *Ostwalds Klassikern* 第 162 期，G·柯瓦列夫斯基（Kowalewski）编，莱比锡，1908 年。并参阅莱布尼兹 *Mathematische Sehriften*，K·J·格哈特（K. J. Gerhardt）编，1849 年起出版。

② *Observatio ...; et de vero sensu methodi infinitesimatis*，第 167—169 页。

斯·德·洛必达（Marquis de L'Hospital）著 *Analyse des Infiniment Petits pour l'Intelligence des courbes*[①]。马奎斯是 J·伯努利的学生，而其师伯努利又以惊人的速度吸收了莱布尼兹的新思想，自己发表了第一本积分教科书[②]。这两本书都代表着近似数学的观点。例如，曲线被看做是短边多边形，切线被看做是这些边的延长。在德国，广泛传播莱布尼兹微分学的是霍尔的克里斯琴·沃尔夫（Christian Wolff），他在 *Elementa matheseos universal*[③] 中发表了他的讲稿的内容。他在微分学一开始就引进了莱布尼兹的微分，尽管他特别强调，微分没有任何实在的等价物。确实，为了有助于我们的直觉，他以富有近似数学味道的方式发展了有关无穷小的观点。他举例说，对于实际测量的目的来说，增添或除去一粒尘土，并不会显著地改变大山的高度。

(4) 你们也会常常看到这样的形而上学观点，认为真正存在着微分。这种观点总是有支持者的，特别是在哲学界，但也有数学物理学家。最重要人物之一是泊松（S. D. Poisson）。他在其名著 *Traité de Mecanique*[④] 中，强烈地表达了这样的观点：无穷小量不仅是一个研究工具，而且是完全的真实存在。

(5) 或许是由于哲学传统，这个观点进入了教科书，甚至到今天还起着显著作用。作为一个例子，我要提到卢勃森（Lübson）著的 *Einleitung in die Infinitesimalrechnung*[⑤] 一书，该书首版于 1855 年，长期以来在一大部分读者中具有非常大的影响。在我那个时代，每个人手上都必有一本卢勃森的书，无论是当中学生的时候或是在

---

① 巴黎，1696 年；第二版，1715 年。

② 译本发表于 *Ostwalds Klassikern* 第 194 期，柯瓦列夫斯基编。伯努利的微分学不久以前由 P·谢弗里特林（P. Schafheitlin）发现，并进行了讨论，见 *Verhandlungen der Naturforscher-Gesellschaft in Basel* 第 32 卷，1921 年。

③ 首版于 1710 年，1742 年马德堡再版，第 545 页。

④ 第二版第一部分第 14 页，巴黎，1833 年。

⑤ 莱比锡第八版，1899 年。

以后的年代，而且许多人是首先受到这本书的激励才去从事进一步的数学研究。卢勃森首先通过极限概念来定义微商，但与此同时，他（在第二版后）把他认为的真正无穷小演算置于用无穷小量进行运算的神秘的章节。这几章都用星号标出，以表示没有得出任何新的结果。微分量不断取有限量直至终极部分而引入的，作为无穷小的、不可指定的量。这些终极部分“虽然不同于绝对零，但是不可指定的、无穷小的量，是一间隙、一瞬间”。英文的引文：“无穷小量是精灵般的转瞬即逝量”（上书第59页、第60页）即由此而来。卢勃森在另一处（第70页）又说：“正如你们所看到的，无穷小方法是非常微妙的，但是正确的。如果不是随逝者而逝、随来者而来，那就仅仅是讲解不当之过。”书中这些段落读起来当然非常有趣。

作为上书的姐妹篇，我再向你们提出广泛采用的沃尔纳（Wüllner）著 *Lehrbuch der Experimentalphysik* 一书的第六版[①]。该书第一卷为理科或医科学生的方便对他们在中学里没有掌握的无穷小演算作了简要的初步讲解，即讲解了为学习物理所不可缺的演算知识。沃尔纳先（从第31页起）讲解了无穷小量 $\mathrm{d}x$ 的意义，然后讲解了第二微分 d^2x，这当然要难一些。我鼓励你们用数学家的眼光去读一读这个导论，并反思一下：由于太难而在中学里压着不教无穷小演算，期望大学第一学期学生从第10页叙述中理解这种演算，是不是有点荒谬？再说那10页不但远远不够，而且太难读了！

为什么与数学上严格的极限方法一起长期以来存在着这样的反思呢？其原因也许必须从一种早已感到的需要中去找，即必须透过对极限方法所作的抽象的逻辑描述，深入到连续量的内在本质，不单纯强调决定极限概念的心理因素，从而对连续量形成较为确定的印象。有一种关于极限的典型描述，以前在教科书和讲课中常常用到，我相信这种说法源自哲学家黑格尔（Hegel）。他说，“函数 $y=f(x)$

---

① 沃尔纳，莱比锡，1907年。

表示事物的存在，而微商表示事物的变化”。这种说法无疑有某些深刻的含义，但是人们必须清楚地认识到，这类话不会促进数学的进一步发展，因为数学的发展必须置于精确的概念之上。

在最近的数学著作里，“实际的”无穷小量再次出现了，但出现在完全不同的领域，如意大利维罗纳派的几何研究，以及希尔伯特的 *Grundlagen der Geometiie*（《几何基础》）[①]。这些研究的主导思想可以简述如下：$x = a$（a 是普通实数）确定的不仅是在 x 轴上的一个点，而是无穷多个点，它们的横坐标之间相差不同阶无穷小量 η, ξ…的有限倍数。只有指定了

$$x = a + b\eta + c\xi + \cdots$$

之后，一个点才确定，其中 a, b, c, …是普通实数，η, ξ…是逐次高阶的实在的无穷小量。希尔伯特使这些新的量 η, ξ…服从某些公理化假设，并使得人们能协调地对它们进行运算。为此目的，最主要的是适当地确定 x 和第二个量 $x_1 = a_1 + b_1\eta + c_1\xi + \cdots$ 之间的关系。第一个假设是：$x >$ 或 $< x_1$ 取决于 $a >$ 或 $< a_1$，或如 $a = a_1$ 则取决于 $b >$ 或 $< b_1$，若更有 $b = b_1$ 则取决于 $c >$ 或 $< c_1$，等等。如果不把这些字母和任何具体的表示联系起来的话，你们就会更清楚地了解这些假设。

如果对这些新的量加上这一条规则，再加上某些其他规则，就能像对待有限量一样对它们进行运算。但是，一个对普通实数系统成立的基本定理现在不再成立：“对给定两正数 e, a，不论 e 如何小，a 如何大，总能找到一个有限整数 n，使 $ne > a$。”事实上，从上面的定义立即可推出，η 的任何有限倍数 $n \cdot \eta$ 总是小于任何正有限数 a。正是由于这个性质，η 才是一个无穷小量。同样，$n \cdot \xi < \eta$，即 ξ 是比 η 高阶的无穷小量。

---

① 第五版，莱比锡，1922 年。

这套数的系统称为非阿基米德系统。上面所述关于有限数的定理称为阿基米德公理，因为阿基米德强调，就他所用的数系而论，这是一个不可证明的假设，或不需证明的基本性质。否定这个公理，就表示可能存在实在的无穷小量。然而，阿基米德公理这个名称，如同许多以个人名字命名的东西那样，不符合历史的真实。在阿基米德之前半个多世纪，欧几里得就认识到这个公理的重要性，不过据说也不是欧几里得发明的，而是由克尼多斯(Knidos)的欧多克斯(Eudoxus)那里来的。欧几里得的许多定理都是这样从别人处来的。对这些在建立非阿基米德几何中作为坐标的非阿基米德量[①]的研究，目的在于加深对连续性的本质的认识。关于普通几何与算术的各种公理之间逻辑依赖性有许多研究，这个研究就属于这一范畴。为此目的，所用方法总是人为地建立一些数系，使原来的公理只有一部分保持成立，从而论证剩下的公理对这些公理的逻辑独立性。

自然会提出这样一个问题：从这样的数系出发，是否可能改变无穷小演算的传统基础，把实在无穷小量包括进去，使之满足现代数学分析的严格要求？换句话说，就是建立起非阿基米德分析。这种分析的首要的问题，是从所设公理出发，证明中值定理

$$f(x+h)-f(x)=h\cdot f'(x+\theta h)。$$

我不愿说不可能沿着这个方向取得进展，但事实是，没有一个埋头研究实在无穷小量的人取得过任何积极的成果。

我要指出，从柯西时代起，教科书里用的“无穷小”这个术语，意义就有若干变化。我们从来不说一个量是无穷地小，而只说它变得无穷地小。这只是表达量无限制地减小到零的一种方便的说法。

我们必须记住在无穷小演算中使用无穷小量所引起的反应。人们不久就意识到这些思想中的神秘、不可证明感，且常常产生一种偏见，认为微分学是一个特殊的哲学系统，它不可证明，只可信奉，或者

---

① 即欧几里得已知的喇叭角，这是非阿基米德量的例子。

直截了当地说，是一个骗局。在这个意义上的最热烈的批评者之一，是哲学家贝克莱(G. Berkeley)主教，他在一本小书 *The Analyst*(《分析者》，伦敦，1734 年)中以滑稽的方式指责了他那个时代数学中所充斥的缺乏明晰性的情况。他声称他有同样的自由来批评数学家据以批评宗教神秘性的种种数学原理及方法，于是对一切新的分析的方法发动了猛烈的攻击，包括用流数以及用微分进行的运算。他得出结论说，整个分析的结构是模糊不清的，而且是完全莫名其妙的。

甚至直到现在，仍然有人持有类似的观点，特别是在哲学界。这或许是由于他们只了解微分运算，而对最近发展起来的严格的极限方法还不全面了解之故。作为一个例子，请让我引用 19 世纪 60 年代出版的、鲍曼(Baumann)著的 *Raum, Zeit und Mathematic* 中的一段话："因而我们放弃莱布尼兹赋予微分运算的逻辑及形而上学的依据，但是不愿触及微分运算本身。在我们看来，它是被实践证明有用的巧妙的发明，是一种技巧而不是一门科学。它不能按逻辑建立起来。它不是从普通数学的成分推导出来的……"[1]

这种反对微分学的反应，从拉格朗日 1797 年发表的《解析函数论》里也可以看出。他不仅想从理论中清除无穷小量，而且想清除一切极限过程。他只讨论能用幂级数表达的函数

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots,$$

而形式地定义"导函数 $f'(x)$"$\left(\text{他很有特色地避免使用微商这个词和符号}\dfrac{\mathrm{d}y}{\mathrm{d}x}\right)$为新的幂级数

$$f'(x) = a_1 + 2a_2x + 3a_3x^2 + \cdots。$$

就是说，他用"导数学"来代替"微分学"。

当然，这种办法并不总是行得通。首先，如我们曾指出的，这里

---

① 鲍曼，第 2 卷，第 55 页，柏林，1869 年。

用的函数的概念太狭窄。此外，这种形式的定义，使我们不可能对导数的性质有更深入的理解。而且，这种定义完全不考虑我们说的“心理上的活动”。对于为什么要对用这种特殊的方法求得的级数感兴趣，完全未作解释。最后，只有完全不考虑级数的收敛性，不考虑级数在什么误差范围内可以用有限项的和来代替的问题，才能一点也不想到极限过程。可是这些问题对于上述级数的实际应用，当然是很重要的。一考虑这些关键问题，就必然要求助于极限概念。然而避免这个概念，又正是发明这个体系的目的。

关于微积分的基础，有不同的意见。说说这些意见的不同处。也许是合适的，因为即使现在，意见的分歧也已超出专业数学家的小圈子。我认为，我们往往能从中找到理解的初步条件，这就又要讲到我们在讲算术基础时提出的一些非常相似的考虑(见第一部分《整数运算的逻辑基础》第 4 点)。在数学的每一个分支里，人们必须截然区分下述两件事：其一，它的结构的内部逻辑相容性；其二，将它的那些公理化地或者说是任意地表述的概念和定理应用于外部的或内部的感性认识的对象上去。康托提到整数时就对“固有的真实性”和“暂时的真实性”作了划分①。前者是指整数的逻辑定义性，后者是指可应用于具体事物。以无穷小演算而论，第一个问题已用极限的概念(以逻辑上完整的方式建立起来的那些理论)完全解决了。第二个问题完全属于知识论的范畴，数学家的贡献仅在于精确地加以表达，把第一部分同它区分出来并加以解决。没有一种纯数学工作按其本性来说能对直接解决第二部分问题有所贡献(见第一部分整数运算的逻辑基础一节，第 4 点)。在无穷小演算的基础方面，这两个完全不同范畴的问题还没有很明确地区分开来。有关的一切争论，就是在这种不利的条件下艰难地进行着的。事实上，这里的第一部分，即纯数学部分，已完全和数学的其他分支一样确立起来了。困难

① 见《数学年刊》第 21 卷(1883 年)，第 562 页。

在于第二部分，即哲学部分。朝着这个方向进行的那些研究，由于这些考虑而具有重大的价值；但是，那些研究绝对必须依赖关于第一个问题的纯数学研究所取得的精确知识。

最后，我必须对无穷小分析的发展作一简短的历史性回顾以结束这一节。当然，我不得不以强调最重要的指导概念为限。这个简短的历史性回顾自然应当通过对当时的全部文献进行彻底研究而扩大。你们可以在马克斯·西蒙在 1896 年法兰克福自然科学家大会上所作的报告（“Zur Geschichte und Philosophie der Differentialrechnung”）中找到许多有趣的资料。

如果我们最终考察一下中学教学中对待无穷小分析的态度，我们也会看到历史发展过程在那里的一定反映。在早些年代，那些教无穷小分析的学校，对建立在极限方法基础上的严格科学结构绝无清楚的概念，至少表现在教科书上，毫无疑问也同样表现在学校教学中。极限方法至多以含糊的方式冒出来，而无穷小量演算却成了主体，有时拉格朗日意义上的微分运算也成了主体。当然，这种教学方法不仅缺乏严格性，而且难以理解。因此对中学里教无穷小演算最终产生明显的反感，就是显而易见的事了。在 19 世纪 70 年代和 80 年代，德国官方甚至命令在“真正的学院”里也禁止这种教学，更说明对此的反感达到了顶点。

事实上，这并未完全防止（我前面已指出过）在必须用到的地方使用极限方法，人们在中学里只是避免使用这个名称，甚至有时以为是在教别的什么东西。这里，我仅提出 3 个例子，你们当中大多数人在回忆上中学时的情形时会想起这些例子来的。

(a) 通过内接和外切正多边形对圆面积和周长进行众所周知的近似计算，显然只不过是一个积分。古代就已使用过这个方法，尤其是阿基米德用过这个方法。事实上，正因为它是经典的，所以至今仍保留在中学教学中。

(b) 物理教学，特别是力学教学，必然要涉及速度和加速度概

念。在包括自由落体定律在内的各种推导中，也一定要用到这些概念。但这些定律的推导，实质上相当于用函数 $z=\frac{1}{2}gt^2+at+b$ 来表示微分方程 $z''=g$ 的积分，其中 a, b 是积分常数。由于物理教学涉及压力，中学里必须解决这个问题，因此或多或少要用到严格的积分方法，当然是以掩盖起来的方式。

(c) 在德国北部的许多中学中，要按所谓谢尔巴赫(Schellbach)方法教最大最小值理论。根据这个方法，令

$$\lim_{x\to x_1}\left(\frac{f(x)-f(x_1)}{x-x_1}\right)=0,$$

以便求得函数 $y=f(x)$ 的极值。这正是微分学的方法，只是没有用微商这个词。谢尔巴赫之所以用上面的表达式，仅仅是因为不许在中学里教微分学，但他不想丢掉这个重要的概念。他的学生原封不动地把这种教法接了过来，并用他的名字来称呼这种方法。结果，费马、莱布尼兹和牛顿所采用的方法，都以谢尔巴赫的名义摆在学生的面前。

最后，让我指出对我们这种改革趋势的态度。在目前德国及其他地方，特别是法国，这种改革的声势越来越大，我们希望它将控制以后几十年的数学教学。我们希望，用符号 $y=f(x)$, $\frac{\mathrm{d}y}{\mathrm{d}x}$, $\int y\mathrm{d}x$ 所表达的那些概念，能以这些符号的形式使学生熟悉起来，不是作为一项新的抽象的科目，而是作为整个教学的一个有机部分，从最简单的例子开始，慢慢地前进。例如，在教十四五岁学生时，可以从处理函数 $y=ax+b$ (a, b 为给定的数)和 $y=x^2$ 开始，在坐标纸上画出它们的图像，使斜率和面积的概念慢慢地发展，但必须坚持用具体的例子来说明。在随后的 3 年里，可将这些知识收集到一起，作为一个整体来处理，结果可能使学生完全掌握无穷小演算的初步知识。关键是要对学生讲清楚，他们所学的不是什么神秘的东西，而是任何人都能懂得的简单事情。

这种改革的迫切性在于，它涉及一些非常重要的数学概念，今天每个数学应用领域都完全受到这些概念的控制，没有它们，大学里的所有学习，甚至实验物理里的最简单的学习，也会成为无本之木。我就只作这些提示，主要是因为这个问题已经在我和席马克合著的那本书里(见本书序)作了充分的讨论。

为了用某些具体内容来补充这些一般的考虑，我现在准备讨论无穷小演算中一个特别重要的主题。

§9.2 泰勒定理

在这里我将用处理三角级数相类似的方法来进行讲解，即从教科书内的一般处理方法出发，先提出很重要的有限级数，再用图形来促进直观的掌握。这种方法似乎总是比较初等，也容易理解。

我们从是否能用最简单的曲线，在短距离范围内，适当地逼近任意曲线 $y=f(x)$ 这个问题开始。最明显的是，在点 $x=a$ 的邻域内，用曲线的切线

$$y=A+Bx$$

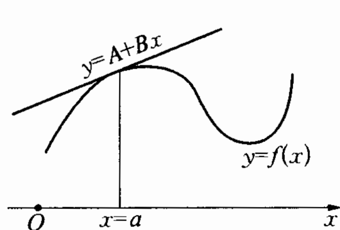


图 9.8

来代替曲线。正像在物理学和其他应用领域中常常舍去在级数展开式中独立变量的高次幂一样(图 9.8)。类似地，通过使用二次、三次…抛物线

$$y=A+Bx+Cx^2,$$

$$y=A+Bx+Cx^2+Dx^3,\ \cdots$$

或用分析术语说，用高次多项式来取得较好的逼近。使用多项式是特别适当的，因为便于计算。我们将给所有这些曲线一个特殊位置，以使它们在点 $x=a$ 处尽可能靠近曲线，即使它们成为密切抛物线。

因此二次抛物线不仅在纵坐标，而且在一阶、二阶导数上均和 $y = f(x)$ 一致（即所谓“密切”）。简单的计算即可表明，n 阶密切抛物线的解析表达式为

$$y = f(a) + \frac{f'(a)}{1}(x-a) + \frac{f''(a)}{1 \cdot 2}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{1 \cdot 2 \cdots n}(x-a)^n \quad (n = 1, 2, \cdots),$$

而这些正是泰勒级数的前 $n+1$ 项和。

这些多项式是否代表有用的近似曲线以及近似到何种程度，我们将用类似于在三角级数中所用的某些试验方法开始研究。我给你们看看几个简单曲线的密切抛物线的图形，这些图形是席马克画的[①]。首先是对下列 4 个在 $x = -1$ 处都有奇点的函数，画出了它们在 $x = 0$ 处的密切抛物线（图 9.9—图 9.12）。

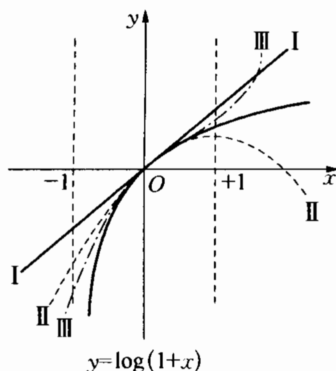


图 9.9

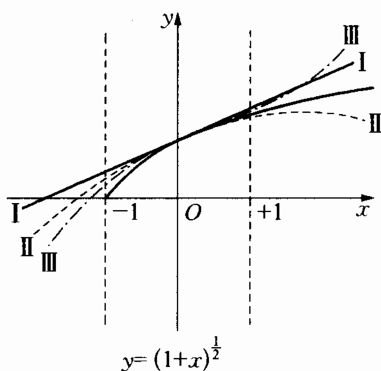


图 9.10

① 这些图中有 4 幅是配合席马克于 1908 年复活节哥廷根大学假期讲座上所作的报告的，报告名为“Über die Gestaltung des mathematischen Unterrichts im Sinne der neueren Reformideen”，刊于 *Zeitschrift für den Mathematischen und naturwissenschaftlichen Unterricht*，第 39 卷（1908 年），第 513 页。也有单印本，莱比锡，1908 年。

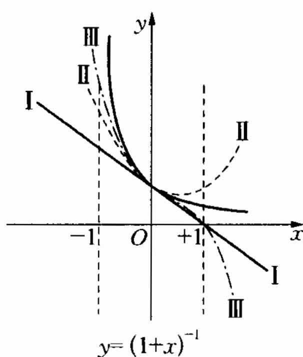


图 9.11

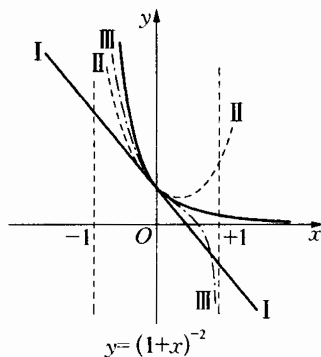


图 9.12

(1) $\log(1+x)\approx x-\frac{x^2}{2}+\frac{x^3}{3}-\cdots$。

(2) $(1+x)^{\frac{1}{2}}\approx 1+\frac{x}{2}-\frac{x^2}{8}+\frac{x^3}{16}-\cdots$。

(3) $(1+x)^{-1}\approx 1-x+x^2-x^3+\cdots$。

(4) $(1+x)^{-2}\approx 1-2x+3x^2-4x^3+\cdots$。

在区间$(-1,\ 1)$中，随着阶数增加，抛物线越来越逼近原来的曲线；但在$x=+1$的右边，它们时而在曲线的上方，时而在曲线的下方，令人注目地越来越离开原来的曲线。

在奇点$x=-1$处，(1)①，(3)，(4)中的原来函数变成无穷，逐次抛物线的纵坐标也不断增大；在(2)中，图中所画这支原曲线在$x=-1$处终止于一条垂直的切线，而所有抛物线都可延伸并超过这一点，但在$x=-1$处变得越来越陡而逼近于原曲线(见图 9.9—图 9.12)。在$x=-1$的对称点$x=+1$处，在前两种情形下，抛物线越来越靠近原曲线。在(3)中，它的纵坐标交替等于 1 和 0，而原曲线的值为$\frac{1}{2}$。最后，在(4)中，纵坐标随阶数增至无穷，并逐次改变符号。

① 此处 1 应除外。——中译者

现在看看两个整超越函数的各个密切抛物线的图形(图 9.13、图 9.14)。

(5) $e^x \approx 1+\frac{x}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}+\cdots$。

(6) $\sin x \approx x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\cdots$。

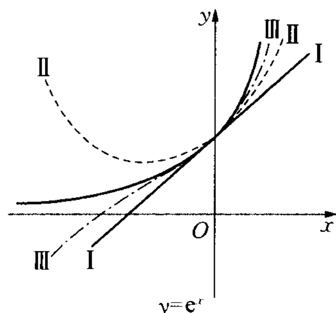


图 9.13

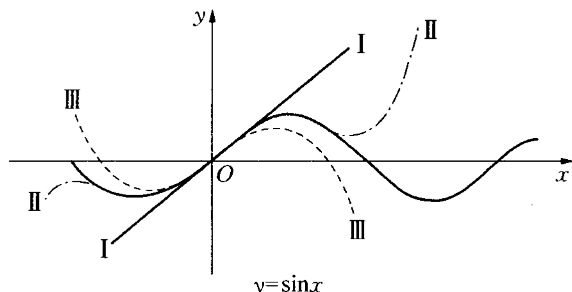


图 9.14

你们注意到,随着它们阶数增加,在越来越大的区间上,抛物线给出原来曲线的有用的近似。特别值得注意的是,在 $\sin x$ 的情形下,抛物线与正弦曲线一样具有越来越多的振荡。

我提醒你们注意,画这些简单情况的图形,也许甚至在中学里也是合适的教材。在有了这些经验的素材后,必须从数学角度来考虑它。第一个问题在于探讨 n 阶密切抛物线与原曲线接近的程度。在实际应用中,这个问题至关重要。这意味着要对纵坐标估计余项,而与此自然相关联的是考察 n 趋向于无穷的过程。这条曲线能否精确地用无穷幂级数来表示。

只需要关于余项

$$R_n(x) = f(x) - \left\{ f(a) + \frac{x-a}{1!} f'(a) + \cdots + \frac{(x-a)^{n-1}}{(n-1)!} f^{(n-1)}(a) \right\}$$

提出一个最常见的定理就足够了。所有的书都给出了此定理的证明，稍后，我将回头来从较一般的观点讨论这个定理。这个定理是：在 a 和 x 之间存在一个 ξ，使 R_n 能表达成

$$R_n(x) = \frac{(x-a)^n}{n!} f^{(n)}(\xi) \quad (a < \xi < x),$$

于是关于判断能否化为无穷级数的问题，归结为当 n 变成无穷时，R_n 是否以 0 为极限。

回到我们的例子，你们可以通过计算证明，在情况(5)和情况(6)下，无穷级数对所有 x 的值都收敛。在情况(1)—(4)中，级数在 -1 与 $+1$ 之间收敛于原来函数，但在此区间之外，级数是发散的。对 $x=-1$，情况(2)的级数收敛于函数值，在(1)，(3)，(4)情况下，级数和函数的极限值都是无穷，所以可以说在此也收敛。但是当级数具有确定的无限极限时仍用收敛这种说法，是不合习惯的。最后考虑在 $x=+1$ 处的情况，这时前两例的级数收敛，后两例中的级数发散。所有这些完全和我们的图像一致。

现在可以考虑一个类似于我们在三角级数的讨论中曾经讨论过的问题，即关于逼近抛物线作为一个完整的曲线的极限位置问题。它们在 $x=\pm 1$ 处当然不可能突然断开。对 $\log(1+x)$ 的情形，我为你们画出了极限曲线(图 9.15)。奇阶和偶阶抛物线有不同的极限位置(图中分别用点线和破折线表示)，它们就是在 -1 和 $+1$ 之间的对数曲线以及垂直线 $x=+1$ 的相应上半部分和下半部分。其他 3 种情况类似。

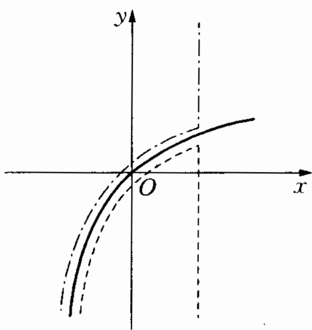


图 9.15

不转入复变量，泰勒级数的理论

考虑就不完整。只有转到复变量上，人们才能理解，为什么在函数是完全解析的地方幂级数却突然不再收敛。事实上，在上述几个例子的情形下，人们可能满足于说，在右方这些级数不可能比在左方更远的地方收敛，而左边由于 $x=-1$ 为奇点，必然不再收敛。但这种理由对下面一类的情况却不适合，函数 $\tan^{-1}x$ 的一支对所有 x 都解析的泰勒展开式

$$\tan^{-1}x \approx x-\frac{x^3}{3}+\frac{x^5}{5}-\cdots$$

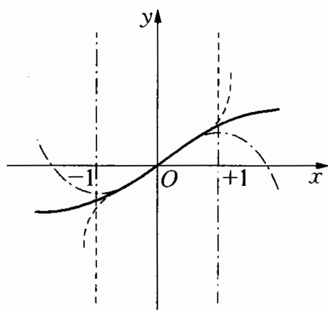


图 9.16

只在区间$(-1,\ +1)$收敛，而密切抛物线交替地收敛于两个不同的极限位置(图 9.16)。在图 9.16 中，第一个极限位置由垂直线 $x=+1$，$x=-1$ 的长点线部分以及位于其间的反正切曲线组成。第二个极限位置，可从前者以垂线的点线部分取代长点线部分而得到。当我们取级数的奇数项时收敛于第一个极限曲线，取偶数项时收敛于第二个。此图中，长点曲线代表 $y=x-\frac{x^3}{3}+\frac{x^5}{5}$，短点曲线代表 $y=x-\frac{x^3}{3}$。如果限于 x 的实数值并注意函数的性态，在完全正则的点$x=\pm 1$ 处突然停止收敛性是无法理解的。可以用收敛圆定理解释这种现象，这是柯西最漂亮的函数论成果之一。该定理说：如果解析函数 $f(x)$是单值时，把它的奇异点标在复平面上，如果 $f(x)$是多值时，则标在黎曼曲面上，这时对应一个正则点 $x=a$ 的泰勒级数，以 a 为圆心而在内部没有奇点的最大圆内收敛(即至少有一个奇点在此圆周上)。在圆外的任何点，级数都不收敛(图 9.17)。

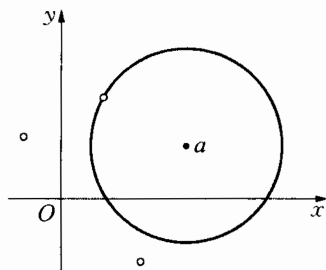


图 9.17

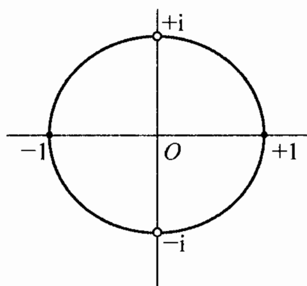


图 9.18

如你们所知，我们的例子 $\tan^{-1}x$ 有奇点 $x=\pm\mathrm{i}$，于是，按 x 的幂展开的级数的收敛圆，是围绕 $x=0$ 的单位圆。因此在 $x=\pm1$ 处收敛性必然停止，因为从此以后实轴上的点超出了收敛圆(图 9.18)。

最后，关于级数在圆周上的收敛性，请你们参考我们在谈及幂级数和三角级数的联系时所作的论述。其收敛性依赖于函数的实部和虚部能否在必然存在奇异点的圆周上展开成收敛的三角级数。

我现在指出泰勒定理与内插问题和有限差分问题的联系，以便使泰勒定理的讨论进行得更生动活泼。这里同样涉及用抛物线逼近给定的曲线问题，但我们不再力图使抛物线在一点附近尽可能地适合，而是使它在许多指定的点与曲线相交。问题仍然是这个内插抛物线给出的近似程度如何？在最简单的情形下，它归结于用割线而不是切线来代替曲线(图 9.19)。类似地，可使二次抛物线通过给定曲线的 3 点，使三次抛物线通过 4 点，等等。

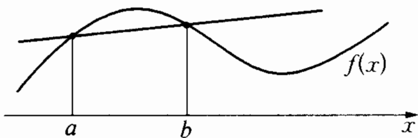


图 9.19

这是一种处理内插问题的自然的方法，它是经常用到的，例如在使用对数表时。在那种情况下，我们是假定对数曲线在两个给定的表值之间沿直线变化，并以众所周知的方法作“线性”插值，并借助差值表来简化这种插值。如果这种近似不够精确，我们就使用二次插值。

对这个一般问题从广泛的角度说明之后，只要令内插抛物线的所有交点重合，我们就可以将泰勒定理中的密切抛物线作为

一个特殊情形。肯定地说，把用密切抛物线代替曲线说成“内插”是不恰当的，除非把“外推”也包含在“内插”问题之内。例如，曲线不仅要与交点之间的割线相比较，也要与交点之外的部分相比较。从整个过程来理解，用逼近这个全面的词似乎比较合适。

现在我将给出最重要的内插公式。我们首先来确定与所给函数在点 $x = a_1, a_2, \cdots, a_n$ 相交，即它是在这些点的纵坐标为 $f(a_1)$，$f(a_2)$，$\cdots$，$f(a_n)$ 的 $n-1$ 次抛物线(图 9.20)。如你们所知，这个问题可用拉格朗日内插公式

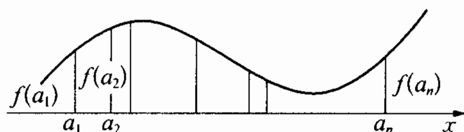


图 9.20

$$
\left.\begin{aligned}
y = &\frac{(x-a_2)(x-a_3)\cdots(x-a_n)}{(a_1-a_2)(a_1-a_3)\cdots(a_1-a_n)}\cdot f(a_1)\\
&+\frac{(x-a_1)(x-a_3)\cdots(x-a_n)}{(a_2-a_1)(a_2-a_3)\cdots(a_2-a_n)}\cdot f(a_2)\\
&+\cdots
\end{aligned}\right\}\tag{1}
$$

解决，这个公式包含具有因子 $f(a_1)$，$f(a_2)$，$\cdots$，$f(a_n)$ 的 n 项。分子逐项地缺少因式 $(x-a_1)$，$(x-a_2)$，$\cdots$，$(x-a_n)$。很容易验证此公式的正确性。因为组成 y 的每一项都是 x 的 $n-1$ 次多项式，从而 y 也是 x 的 $n-1$ 次多项式。如令 $x = a_1$，则除第一项分式化成 1 外，其余各项均为 0，故有 $y = f(a_1)$。类似地，对 $x = a_2$ 有 $y = f(a_2)$，等等。

由这个公式，通过特殊化，很容易推出常常被称为牛顿公式的结果，即横坐标 $a_1, \cdots, a_n$ 为等距的情况(图 9.21)。由于在这种情况下使用有限差分计算符号更为有利，我们首先来介绍它。

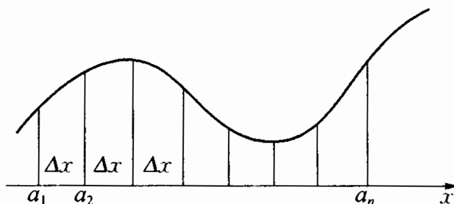


图 9.21

设 Δx 为 x 的任意增量,令 $\Delta f(x)$ 表示 $f(x)$ 的对应增量,因而

$$f(x+\Delta x)=f(x)+\Delta f(x),$$

因为 $\Delta f(x)$ 也是 x 的函数,当我们再给 x 以增量 Δx 时,它又有有限差分,称为二阶差分 $\Delta^2 f(x)$,即有

$$\Delta f(x+\Delta x)=\Delta f(x)+\Delta^2 f(x)。$$

同样地,有

$$\Delta^2 f(x+\Delta x)=\Delta^2 f(x)+\Delta^3 f(x)$$

等等。这个符号与微分学中的符号完全相同,只不过这里讲的是有限量,而不存在过渡到极限的问题。

从上面的差分定义立即可以推出 $f(x)$ 在相继的等距离处之值:

$$\begin{cases} f(x+\Delta x)=f(x)+\Delta f(x), \\ f(x+2\Delta x)=f(x+\Delta x)+\Delta f(x+\Delta x) \\ \qquad\qquad\quad =f(x)+2\Delta f(x)+\Delta^2 f(x), \\ f(x+3\Delta x)=f(x+2\Delta x)+\Delta f(x+2\Delta x) \\ \qquad\qquad\quad =f(x)+3\Delta f(x)+3\Delta^2 f(x)+\Delta^3 f(x), \\ f(x+4\Delta x)=f(x)+4\Delta f(x)+6\Delta^2 f(x)+4\Delta^3 f(x) \\ \qquad\qquad\qquad +\Delta^4 f(x)。 \end{cases} \tag{2}$$

这个表还可延续下去,在各等距点的值,由在初始点的逐次差分附上二项式系数作为因子来表达。

对 x 轴上 n 个等距点

$$a_1 = a,\ a_2 = a + \Delta x,\ \cdots,\ a_n = a + (n-1)\Delta x$$

的 $(n-1)$ 阶内插抛物线的牛顿公式，即在这些点与 $f(x)$ 的纵坐标相同的内插抛物线的表达式为

$$y = f(a) + \frac{(x-a)}{1!}\cdot\frac{\Delta f(a)}{\Delta x} + \frac{(x-a)(x-a-\Delta x)}{2!}\cdot\frac{\Delta^2 f(a)}{(\Delta x)^2} + \cdots + \frac{(x-a)(x-a-\Delta x)\cdots(x-a-(n-2)\Delta x)}{(n-1)!}\cdot\frac{\Delta^{n-1} f(a)}{(\Delta x)^{n-1}}, \tag{3}$$

这是 x 的一个 $n-1$ 次多项式。当 $x=a$ 时化为 $f(a)$，当 $x=a+\Delta x$ 时，除前两项外，其余均变为零，因而剩下 $y=f(a)+\Delta f(a)$，而根据(2)式，它等于 $f(a+\Delta x)$，等等。因此(3)式产生了在所有 n 个位置都取正确值的多项式。

如果想实际应用这个内插多项式，必须知道用它来表示 $f(x)$ 的正确性，即必须有可能去估计余项。柯西在 1840 年给出了这个公式①，我愿在此推导如下。从更一般的拉格朗日公式出发，设 x 是 $a_1,\ a_2,\ \cdots,\ a_n$ 之间或之外的任何值（内插或外推）。用 $P(x)$ 表示由公式给出的内插多项式，用 $R(x)$ 表示余项

$$f(x) = P(x) + R(x)。\tag{4}$$

根据 $P(x)$ 的定义，余项 R 在 $x = a_1,\ a_2,\ \cdots,\ a_n$ 处为零，因此有

$$R(x) = \frac{(x-a_1)(x-a_2)\cdots(x-a_n)}{n!}\psi(x)。$$

把因子 $n!$ 提出来是很方便的，于是，与泰勒级数余项完全类似，可

① *Comptes Rendus*，第 11 卷，第 775—789 页。*Œuvres*，第 5 卷第 1 分册，第 409—424 页，巴黎，1885 年。

知 $\psi(x)$等于 $f(x)$的 n 阶导数在$n+1$ 个点 a_1, a_2, …, a_n, x之间的某处 $x=\xi$之值。如果考虑到当 $f^{(n)}(x)=0$ 时,$f(x)$ 是一个多项式,那么关于 $f(x)$和 $n-1$ 阶多项式的偏离程度依赖于函数 $f^{(n)}(x)$ 的全部性质,这一论断就有说服力了。

关于余项公式的证明,我们推导如下。令下式为新变量 z 的函数:

$$F(z)=f(z)-P(z)-\frac{(z-a_1)(z-a_2)\cdots(z-a_n)}{n!}\psi(x),$$

其中 x 为在 $\psi(x)$内的一个参量。因为 $P(a_1)=f(a_1)$, $P(a_2)=f(a_2)$, …, $P(a_n)=f(a_n)$, 所以, $F(a_1)=F(a_2)=\cdots=F(a_n)=0$。因为当 $z=x$ 时,最后一项化为 $R(x)$,从而右端此时为 0,故 $F(x)=0$。因此,我们知道 $F(z)$有 $n+1$ 个零点 $z=a_1$, a_2, …, a_n, x。现在对其应用推广的中值定理,这是通过反复使用普通中值定理而得出的,即若具有 n 阶导数的连续函数在 $n+1$ 个点为零,则在包含所有零点的区间内至少有一点,使其 n 阶导数为零。因此,如果 $f(z)$,从而 $F(z)$有 n 阶连续导数,则在 a_1, a_2, …, a_n, x 的两端值之间必然存在值 ξ,使得

$$F^{(n)}(\xi)=0。$$

但我们有

$$F^{(n)}(z)=f^{(n)}(z)-\psi(x),$$

因为$n-1$ 次多项式 $P(z)$的 n 阶导数为 0,而最后一项只有最高次项 $\dfrac{z^n\psi(x)}{n!}$的 n 阶导数不为零。因此,最后得

$$F^{(n)}(\xi)=f^{(n)}(\xi)-\psi(x)=0 \text{ 或 } \psi(x)=f^{(n)}(\xi)。$$

此即所要证明的。

我们将带余项的牛顿内插公式表达如下

$$f(x) \doteq f(a) + \frac{x-a}{1!} \cdot \frac{\Delta f(a)}{\Delta x} + \frac{(x-a)(x-a-\Delta x)}{2!} \cdot \frac{\Delta^2 f(a)}{(\Delta x)^2} + \cdots + \frac{(x-a)\cdots[x-a-(n-2)\Delta x]}{(n-1)!} \cdot \frac{\Delta^{(n-1)} f(a)}{(\Delta x)^{n-1}} + \frac{(x-a)\cdots[x-a-(n-1)\Delta x]}{n!} f^{(n)}(\xi), \tag{5}$$

其中 ξ 是包含 $n+1$ ①个点 a，$a+\Delta x$，$a+2\Delta x$，…，$a+(n-1)\Delta x$，x 的区间内的某个值。事实上，公式(5)在应用中是不可缺少的。我们已提到使用对数表时用到线性插值。设 $f(x)=\log x$，且 $n=2$，从式(5)可得

$$\log x = \log a + \frac{x-a}{1!} \cdot \frac{\Delta \log a}{\Delta x} - \frac{(x-a)(x-a-\Delta x)}{2!} \frac{M}{\xi^2}。$$

由于 $d^2 \frac{\log x}{dx^2} = \frac{-M}{x^2}$，其中 M 是对数系的模数，因此，我们有了在对数表里对 a 与 $a+\Delta x$ 之间作线性插值时可能产生的误差的表达式。此误差的符号按 x 在 a 与 $a+\Delta x$ 之间或在此区间之外而不同。每个需要用对数表的人都应该真正学会这个公式。

我不打算在应用上多说些什么，只提醒你们注意在牛顿内插公式和泰勒公式之间的显著类似之处。这种类似有一个基本原因。从牛顿公式很容易精确推出泰勒定理，这相应于内插多项式求极限后，就变为密切多项式。因此，如果保持 x，a 和 n 固定，而令 Δx 趋向于0，则因为 $f(x)$ 有 n 阶导数，(5)式中的 $n-1$ 个差商变成各阶导数：

$$\lim_{\Delta x \to 0} \frac{\Delta f(a)}{\Delta x} = f'(a), \lim_{\Delta x \to 0} \frac{\Delta^2 f(a)}{\Delta x^2} = f''(a), \cdots。$$

在(5)式的最后一项中，随 Δx 的减小，ξ 可能改变。但是，由于右端

① 英文原文误为 $n-1$。——中译者

所有其他项均有确定的极限，而在极限过程中，左端有固定值 $f(x)$，故 $f^{(n)}(\xi)$ 必然收敛于一个确定值。又因 $f^{(n)}(x)$ 的连续性，所以此值必然是函数在 a 与 x 之间某处之值。如果仍用 ξ 表示这个位置，则得

$$f(x)=f(a)+\frac{x-a}{1!}f'(a)+\cdots+\frac{(x-a)^{n-1}}{(n-1)!}f^{(n-1)}(a)$$
$$+\frac{(x-a)^n}{n!}f^{(n)}(\xi)\quad(a<\xi<x)。$$

因此，对带余项的泰勒定理得到了一个完全的证明，同时使它在插值理论中占据了一个合理的位置。

我个人认为，这种证明法是泰勒定理最好的求证法，它使泰勒定理与一些十分简单的问题广泛地联系起来，并提出了一个过渡到极限的自然途径。但熟知这些事情的数学家并不这么想(值得注意的是，许多人，甚至包括许多教科书的作者，却不了解以上道理)。他们习惯于板起面孔对待过渡到极限的问题，宁愿直接证明泰勒定理而不把它和有限差分计算联系起来。

然而，作为一个历史事实，我必须强调泰勒定理确实是起源于有限差分计算的。我曾说过，布洛克·泰勒在他的 *Methodus incrementorum*(《增量方法》)中首先发表了此定理[①]。他首先推导了牛顿公式，当然是不带余项的，然后令 $\Delta x=0$ 和 $n=\infty$，从而由牛顿公式的前面若干项正确地得到他的新级数的前面若干项：

$$f(x)=f(a)+\frac{x-a}{1!}\cdot\frac{\mathrm{d}f(a)}{\mathrm{d}a}+\frac{(x-a)^2}{2!}\cdot\frac{\mathrm{d}^2f(a)}{\mathrm{d}a^2}+\cdots。$$

按同样法则把这个级数推导下去，对他来说似乎是不言而喻的，他既未考虑过余项，也未考虑过收敛性。由此过渡到极限，实在是大胆之极。前面一些项中只出现 $x-a-\Delta x$，$x-a-2\Delta x$，… 不会有什么

---

① 伦敦，1715年，第21—23页。

困难，因为 Δx 的有限倍数随 Δx 一起趋向于零。但随着 n 的增加，项数也越来越多，从而给出了 k 越来越大的因式 $x-a-k\Delta x$，没有理由仍用同样方法处理这种情况并假定它们变成一个收敛级数。

这里，泰勒和莱布尼兹学派的人一样，实际上是在毫不怀疑的情况下作了无穷小量（微分）运算。有趣的是，尽管他在牛顿眼里不过是一个 29 岁的年轻人，但他却脱离了牛顿的极限方法。

关于泰勒定理的全部发展经过，你们可以在艾尔弗雷德·普林斯海姆（Alfred Pringsheim）的专论“Zur Geschichte des Taylorschen Lehrsatzes”[①]中找到极好的批评分析。

我愿意在这里讲讲泰勒级数和麦克劳林级数之间通常的差别。众所周知，许多教科书把 $a=0$ 的泰勒级数的特殊情况

$$f(x)=f(0)+\frac{x}{1!}f'(0)+\frac{x^2}{2!}f''(0)+\cdots$$

称为麦克劳林级数；而且许多人可能认为这个区别是重要的。然而，任何一个了解情况的人都把这个区别相对看成数学上无关紧要的事。但许多人并不了解，从历史上来考虑，这纯粹是误解。因为泰勒毫无疑问已先用上述方法推导了他的一般定理，而且在《增量方法》一书后面的某处（该书第 37 页），他强调了 $a=0$ 时的特殊情形，并指出他可以用现今称为待定系数法的方法直接推出。而后，麦克劳林于 1742 年在他的 *Treatise of Fluxions*（《流数论》）中把这个推导方法拿了过去[②]，他明确地引用了泰勒的推导，并未声称有何创新。但他的引用似乎没有被当作一回事，于是麦克劳林就似乎被当作发明定理的人了。这类张冠李戴的事是常有的。后来正本清源，才把这个一般定理命名为泰勒定理。但是这类根深蒂固的讹误，即使不是不可能纠正，也是很难纠正的，顶多只能在一小圈对掌故有兴趣的

---

① 艾尔弗雷德·普林斯海姆，《数学丛书》第 1 卷（1900 年）第三分册，第 433—479 页。

② 爱丁堡，1742 年，第 2 卷，第 610 页。

人中讲讲实情而已。

下面我要发点一般性的议论，以补充对无穷小演算的讨论。

§9.3 历史的与教育学上的考虑

首先我要提醒你们，泰勒建立的差分法和微分学之间的联系，保持了很长一段时间。这两个分支总是并肩发展的，在欧拉的解析展开里还是如此，微分学公式仍作为差分学里初等关系的极限情形而出现。首先打破这种自然联系的是前面常常提到的拉格朗日导数学的形式定义。我愿向你们介绍一本 18 世纪末拉克罗依克斯(Lacroix)编的一本书，书名为 *Traite du Calcul Dif f'erentiel et du Calcul Integral of Lacroix*。这本书紧接着拉格朗日之后，把当时所有的关于无穷小演算的知识都汇集在一起了①。作为反映该书特色的一个典型例子，让我们来看一下书中(第一卷，第 145 页)的导数定义：一个函数 $f(x)$由一个幂级数来确定。利用二项式定理(重新排列各项)，得

$$f(x+h)=f(x)+hf'(x)+\frac{1}{2}h^2f''(x)+\cdots,$$

著者将幂级数的一次项记为 $\mathrm{d}f(x)$，用 $\mathrm{d}x$ 表示 h，于是他得到他称为微分系数的导数

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x}=f'(x)。$$

由此可知，这个式子的推导方式是非常肤浅的，即使表面上看无懈可击。在这些想法范围内，著者当然不可能把差分法作为出发点。然而由于他认为这个分支在实用中十分重要，不可忽略，因此他采用了独立发展这个分支的手段。他在第三卷中十分彻底地做到了这一

① 拉克罗依克斯，3 卷本，巴黎，1797—1800 年。后又多次重版。

点，但根本没有把这个分支同微分学联系起来。

这个“大本 Lacroix”作为 19 世纪出版的许多无穷小演算教程的依据，有着重要的历史意义。在这些第一流的教材里，我要提到拉克罗依克斯自己写的教科书“小本 Lacroix”[①]。

从 19 世纪 20 年代以来，教科书都强烈地受到柯西的极限方法的影响。柯西把极限方法提到了一个光荣的地位。我们首先想到的是许多法国教科书，其中大多数，例如，*Cours d'analyse de l'Ecole Politechnique* 是大学教材。德国的教材都直接或间接地以法国教材为蓝本，只有施洛米尔赫(Schlömilch)的教本为唯一例外。在长长的书单中，我只挑出塞雷特的 *Cours de Calcul Différentiel et Intégral* 来谈一谈。这本书于 1869 年首版于巴黎，后于 1884 年由阿历克斯·哈奈克(Alex Harnack)译成德文，从此成为我们用得最广的教科书之一。它经过一大批修订者之手，内容不够匀称。但 1906 年后的版本经过沙洛登堡的 G·谢弗斯的彻底修订，成了 G·谢弗斯一手修订的工作[②]。我也很愿意提到一本全新的法文书——古尔萨(E. Goursat)著的 3 卷本 *Cours d'Anlyse Mathématique*[③]，这本书在很多方面比塞雷特的书充实，特别包含着一大堆完全属于现代进展的内容。其次，这是一个很可读的书。

在最近出的所有教材中，导数和积分完全建立在极限概念之上，对差分法或插值问题一句也没有提。或许按这个方法，人们觉得可以把问题看得更清楚些，但从另一个方面来说，视野太窄了些，好像我们在用一台显微镜看东西一样。差分学现在已全部交给了实用的

---

① *Traité Elémentaire du calcal Différentiel et Intégral*，两卷本，巴黎，1797 年。

② 自 1906 年后署名 J·A·塞雷特及 G·谢弗斯合著，书名为 *Lehrbuch der Differentialund Integralrechnung*，第 1 卷出了第六版，莱比锡，1915 年；第 2 卷出了 6—7 版；第 3 卷出了第五版，1914 年。

③ 巴黎，1902—1907 年，第 1 卷第三版，1917 年；第 2 卷第三版，1918 年；第 3 卷第二版，1915 年。第 1 卷由 E·R·赫德里克(E. R. Hedrick)译成英文，出版于 1904 年，吉恩出版股份公司；第 2 卷由 E·R·赫德里克与 O·邓克尔(O. Dunkel)合译，1916 年由吉恩出版股份公司出版。

计算者，特别是天文学家要用到它，但数学家已听不到它了。我们希望将来有所改变①。

作为无穷小演算讨论的总结，我愿意再次强调与教科书里的习惯作法特别不同的我的 4 个观点：

(1) 借助图形(在傅里叶和泰勒级数里用逼近曲线)来说明抽象的思想。

(2) 强调与相邻领域的关系，诸如与差分法、插值的关系，最后是和哲学研究的关系。

(3) 强调历史的发展。

(4) 展出一些普及著作的样本，以说明受到这些著作影响的大众观念与受过专门训练的数学家的观念的差异。

我认为，尤其是未来有前途的教师，应把所有这 4 点都加以考虑，这是极其重要的。你们一开始执教，就会面临那些流行的观点。如果缺乏判别能力，如果没有掌握数学的直观成分以及它与邻近领域的重要关系，如果不了解历史的发展，你们的脚跟就站不稳。那样的话，你们或者会退缩到最现代的纯粹数学的阵地上，使中学生不能接受，或者屈服于攻击，把你们在大学里学的东西都收起来，甚至使你们的教学内容淹没在传统的习惯之中。我常常谈的中学和大学教学的不衔接，正是在无穷小演算这个领域中最严重。我希望我的话有利于克服这种互相脱节的现象，也许能为你们的教学提供一些有用的方法。

我在此结束关于常规分析的讨论。通过附录部分，我将讨论我曾偶然提及的现代数学的若干理论。我想，做为一名教师，对此应有所知。

---

① 为了有一个开端，我促请弗里森多尔夫(Friesendorff)及普卢姆(Prümm)把马尔柯夫斯(Markoffs)的《微分学》译成了德文(莱比锡，1896 年)。在《百科全书》中有一系列文章。新出了一本诺伦德(Nörlund)的《微分学》(柏林，施普林格出版社，1924 年)，这本书是用新观点写的。

附　　录

Ⅰ. 数 e 和 π 的超越性

第一个课题，我将讨论数 e 和 π，特别是我希望证明它们是超越数。

从古代开始就以几何问题的形式对 π 发生了兴趣。甚至在那时已将它的近似计算和它的严格理论构造问题区分开来，并为两个问题的解决打下了一定的基础。在第一个问题方面，阿基米德借助于内接和外切多边形逼近圆的方法取得了实质性的进展。第二个问题不久就集中到是否能用直尺和圆规作出 π 来。试用了所有方法，但没有怀疑过不断失效的原因是由于作图的不可能性。某些早期的尝试已由鲁迪奥（Rudio）[①]汇集发表了。化圆为方的问题仍然是最受关注的问题之一，正如我已提到的，许多人企图自寻出路以求解决这个问题，而不知道或者不相信，现代科学早已回答了这个问题。

事实上，这些古老的问题今天早已完全解决。人们有时倾向于怀疑人类的知识是否真正能向前发展，而在某些领域内，这个怀疑可能是正确的。但在数学上，我有一个例子说明确实不必怀疑。

以上问题现代解法的基础，可以从牛顿和欧拉之间的时期说起。对 π 进行近似计算的有力工具是无穷级数，它可以按所需的准确度进行计算。最精心的结果是英国人向克斯（W. Shanks）获得的结

---

① 鲁迪奥，*Der Bericht des Simplicius über die Quadraturen des Antiphon und Hippokrates*，莱比锡，1908 年。

果，他把 π 计算到了 707 位①。这只好当作像运动员热衷于打破纪录一样的行动，因为任何应用都不会要求这样的准确度。

在理论方面，我们发现在同一时期，开始了对自然对数的底即数 e 的研究。著名的关系式 $e^{i\pi}=-1$ 已被发现，在积分学中已研究出了对于最终解决圆的求积（即化圆为方）的问题至关重要的一种方法。解决此一问题的关键步骤是埃尔米特（Hermite）在 1873 年证明了 e 的超越性②。他未能证出 π 的超越性。这件事是由林德曼在 1882 年完成的③。

这些结果是对上述古典问题的本质推广。以前只关心用圆规直尺作出 π，分析上相当于用有限次开平方根和有理数来表示 π 的问题，但现代的结果不限于证明这种表示的不可能性，而是证明了更深刻的结果，即 π（和 e 一样）是超越数，即它不满足任何整数系数的代数关系。换句话说，e 和 π 都不可能是一个整数系数方程式

$$a_0+a_1x+a_2x^2+\cdots+a_nx^n=0$$

的根，不论 a_0，…，a_n 和次数 n 取得多大。系数是整数这一点是关键的。当然对分数也行，因为乘以公分母可将系数变为整数。

我们现在用希尔伯特在 1893 年《数学年刊》第 43 卷中所给出的简化方法，证明 e 的超越性。我们将指出，如假设存在等式

$$a_0+a_1e+a_2e^2+\cdots+a_ne^n=0 \quad (\text{其中 } a_0\neq 0), \tag{1}$$

其中 a_0，a_1，…，a_n 是整数，将导致矛盾。证明中只用到整数的最简单性质。从数论上说，只用到可除性的最初等的定理，具体来说，用到整数的质因子分解的唯一性，以及质数的个数是无穷的这两点。

---

① 向克斯，参阅韦伯—韦尔斯坦因教科书，第 1 卷，第 523 页。

② 埃尔米特：*Comptes Rendus* 第 77 卷，1873 年，第 18—24，74—79，226—233，285—293 页，或其著作第 3 卷，1912 年，第 150 页。

③ *Sitzungsberichte der Berliner Akademie*，1882 年，第 679 页，以及《数学年刊》第 20 卷，1882 年，第 213 页。

证明的方案如下。我们要建立一个方法,使能用有理数特别好地近似于 e 和 e 的幂,使得我们有

$$e = \frac{M_1 + \varepsilon_1}{M},\ e^2 = \frac{M_2 + \varepsilon_2}{M},\ \cdots,\ e^n = \frac{M_n + \varepsilon_n}{M}, \tag{2}$$

其中 $M, M_1, M_2, M_3, \cdots, M_n$ 是整数, $\varepsilon_1/M, \varepsilon_2/M, \cdots, \varepsilon_n/M$ 是非常小的正的纯小数。于是,在乘以 M 后,方程(1)化成

$$(a_0M + a_1M_1 + a_2M_2 + \cdots + a_nM_n) + (a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_n\varepsilon_n) = 0。 \tag{3}$$

第一个括号内是整数,我们将证明它不是零。至于第二个括号,我们将证明可使 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 充分小,成为一个正的纯小数。于是得出不为零的整数 $a_0M + a_1M_1 + \cdots + a_nM_n$ 加上一个纯小数 $a_1\varepsilon_1 + \cdots + a_n\varepsilon_n$ 成为零的矛盾结论,这就表明(1)式是不可能的。

在刚才我已概述的讨论过程中,我们将用到如果一个整数不能被确定数整除则此整数不可能为零这个定理(因为零可被每个数整除)。也就是说,我们将证明 $M_1, \cdots, M_n$ 可被某质数 P 整除,但 a_0M 没有因子 P,于是 $a_0M + a_1M_1 + \cdots + a_nM_n$ 不可能被 P 整除。因此不等于零。

实现上述证明的主要工具是埃尔米特为此目的而发明的定积分,我们称之为埃尔米特积分。这个证明的关键在于它的结构。这个定积分之值为正整数,并用来定义 M

$$M = \int_0^\infty \frac{z^{p-1}((z-1)(z-2)\cdots(z-n))^p e^{-z}}{(p-1)!} dz, \tag{4}$$

其中 n 是所设方程(1)的次数, p 是一个稍后将确定的质数。将积分 Me^v 的积分区间在点 v 分开,并令

$$M_v = e^v \int_v^\infty \frac{z^{p-1}((z-1)(z-2)\cdots(z-n))^p e^{-z}}{(p-1)!} dz, \tag{4a}$$

$$\varepsilon_v = e^v \int_0^v \frac{z^{p-1}((z-1)(z-2)\cdots(z-n))^p e^{-z}}{(p-1)!} dz。\quad (4b)$$

我们将得到对 $e^v(v=1, 2, \cdots, n)$ 所需的近似式(2)。

现在给出证明的细节。

(1) 我们从 Γ 函数理论开始的著名公式

$$\int_0^\infty z^{\rho-1} e^{-z} dz = \Gamma(\rho)$$

出发。我们只对整数值 ρ 用此公式。此时有 $\Gamma(\rho) = (\rho-1)!$，且将在这个限制下来推导它。如果在 $\rho > 1$ 时对它进行分部积分

$$\int_0^\infty z^{\rho-1} e^{-z} dz = (-z^{\rho-1} e^{-z}) \Big|_0^\infty + \int_0^\infty (\rho-1) z^{\rho-2} e^{-z} dz$$

$$= (\rho-1) \int_0^\infty z^{\rho-2} e^{-z} dz。$$

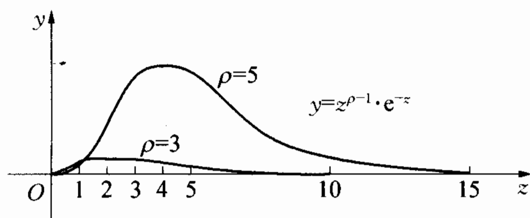
右端的积分除去 z 的指数降低了以外，和左边的形式完全一样。如果反复实施这个过程，因 ρ 是一个整数，最终必然得到 z^0；又因 $\int_0^\infty e^{-z} dz = 1$，故可得

$$\int_0^\infty z^{\rho-1} e^{-z} dz = (\rho-1)(\rho-2)\cdots 3 \cdot 2 \cdot 1 = (\rho-1)!。\quad (5)$$

因此对于整数 ρ，这积分是一个随 ρ 而非常迅速增长的一个正整数。

为了从几何上阐明这个结果，画出不同的 ρ 值的曲线 $y = z^{\rho-1} e^{-z}$。积分值于是由曲线下面伸延至无穷的面积所表示(附图 1)。ρ 越大，曲线在原点越靠拢 z 轴，但在越过 $z=1$ 的地方上升得很快。对所有 ρ 值，曲线在 $z = \rho-1$ 处有最大值。换句话说，随 ρ 增加，最大值发生在越来越右的地方。其值也随 ρ 增加。在最大值的右边，e^{-z} 起决定作用，因此曲线下降，逐渐逼近 z 轴。由此可知，此面积(我们的积分)总保持有限值，但随 ρ 的增加而迅速增大。

(2) 利用这个积分，很容易求出埃尔米特积分的值。用二项式



附图 1

定理展开被积函数

$$((z-1)(z-2)\cdots(z-n))^p=(z^n+\cdots+(-1)^n n!)^n$$
$$=z^{np}+\cdots+(-1)^n(n!)^p,$$

其中只写出 z 的最高和最低次幂的项。积分变成

$$M=\frac{(-1)^n(n!)^p}{(p-1)!}\int_0^\infty z^{p-1}e^{-z}dz+\sum_{p=p+1}^{np+p}\frac{c_\rho}{(p-1)!}\cdot\int_0^\infty z^{\rho-1}e^{-z}dz。$$

c_ρ 是由二项式定理产生的整系数。现在可将公式(5)应用于每一个这样的积分,并求得

$$M=(-1)^n(n!)^p+\sum_{\rho=1+p}^{np+p}c_\rho\frac{(\rho-1)!}{(p-1)!}。$$

和式的下标 ρ 总是大于 p,因而 $(\rho-1)!/(p-1)!$ 是一个含 p 作为因子的整数,所以可以将 p 作为公因子提到整个和式之外

$$M=(-1)^n(n!)^p+p(c_{p+1}+c_{p+2}(p+1)$$
$$+c_{p+3}(p+1)(p+2)+\cdots)。$$

现在,就被 p 的可除性而论,M 的性质将取决于第一项 $(-1)^n(n!)^p$。因为 p 是质数,如果它不能整除 1, 2, …, n,则必然不能整除此项,而当 $p>n$ 时就是这种情况。因为质数有无穷个,所以可以有无限多个方式满足此条件,因而能使 $(-1)^n(n!)^p$,从而也使 M 不能被 p 整除。

再说 $a_0 \neq 0$。按上述，可以选择 p 也大于 $|a_0|$，从而使 a_0 不能被 p 整除。不过这样的话，a_0M 就不能被 p 整除，而这正是我们所要证明的。

(3) 现在必须考察由(4a)式所定义的数 $M_v(v=1, 2, \cdots, n)$。将因式 e^v 放入积分号内，并引入新的积分变量 $\zeta = z - v$，当 z 从 v 变到 ∞ 时，ζ 从 0 变到无穷大。我们有

$$M_v = \int_0^{\infty} \frac{(\zeta+v)^{p-1}((\zeta+v-1)(\zeta+v-2)\cdots(\zeta+v-n))^p e^{-\zeta}}{(p-1)!} d\zeta。$$

这个表达式的形式和前面对 M 的表达式完全相同，可以用同样的方法处理。如果将被积函数的因式乘开，将得到最低幂为 ζ^p 的一组整系数的 ζ 之幂，分子的积分于是成为下列积分的组合

$$\int_0^{\infty} \zeta^p e^{-\zeta} d\zeta, \int_0^{\infty} \zeta^{p+1} e^{-\zeta} d\zeta, \cdots, \int_0^{\infty} \zeta^{(n+1)p-1} e^{-\zeta} d\zeta。$$

而按(5)式，它们分别等于 $p!, (p+1)!, \cdots$，因而分子将是 $p!$ 乘以某个整数 A，故我们有

$$M_v = \frac{p!A_v}{(p-1)!} = PA_v \quad (v=1, 2, \cdots, n)。$$

换句话说，每个 M_v 是一个能被 p 整除的整数。加上前面第二款所述的结果，即可证明 $a_0M + a_1M_1 + \cdots + a_nM_n$ 不能被 p 整除，因此不为零。

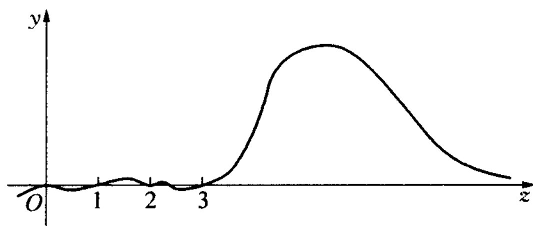
(4) 证明的第二部分涉及和 $a_1\varepsilon_1 + \cdots + a_n\varepsilon_n$，其中，按(4b)式有

$$\varepsilon_v = \int_0^{v} \frac{z^{p-1}((z-1)(z-2)\cdots(z-n))^p e^{-z+v}}{(p-1)!} dz。$$

必须证明，适当选择 p 后，可使 ε_v 变得任意小。为此，要利用可以使 p 任意大这个条件。因为至今加在 p 上的唯一条件是它必须是大于 n 也大于 a_0 的质数，这些条件可以用任意大的质数来满足。

我们来看被积函数的图形。在 $z=0$ 处它与 z 轴相切，但在 $z=$

1, 2, …, n(在附图 2 里, $n=3$), 因为 p 是奇数,它既与 x 轴相切,又与它相交。不久会看到,在分母中出现 $(p-1)!$, 使得 p 很大时,在区间$(0, n)$内曲线仅稍稍偏离 z 轴,因此有理由使得 ε_v 十分小。对 $z>n$, 曲线上升而且渐近地趋向前面的曲线 $z^{\rho-1}e^{-z}(\rho=(n+1)p)$, 并最终逼近于 z 轴。正因如此,积分的值 M(该处是从 0 到∞取积分)随 p 而迅速增加。



附图 2

在实际估计这些积分时,一个粗糙的近似就够用了。设 G 和 g_v 分别是函数 $z\cdot(z-1)\cdots(z-n)$ 和 $(z-1)(z-2)\cdots(z-n)e^{-z+v}$ 在区间$(0, n)$的绝对值的最大值

$$\left.\begin{array}{l}|z(z-1)\cdots(z-n)|\leqslant G\\ |(z-1)(z-2)\cdots(z-n)e^{-z+v}|\leqslant g_v\end{array}\right\}0\leqslant z\leqslant n。$$

因为函数的积分总不会大于其绝对值的积分,故对每个 ε_v 我们有

$$|\varepsilon_v|\leqslant\left\{\int_0^v\frac{G^{p-1}g_v}{(p-1)!}dz=\frac{G^{p-1}g_v\cdot v}{(p-1)!}\right\}。\tag{6}$$

现在,G, g_v 和 v 是与 p 无关的固定数,但分母 $(p-1)!$ 随 p 增长的速度最终要比 G^{p-1} 大得多,或较准确地说,当 p 充分大时,分数 $G^{p-1}/(p-1)!$ 会变得比任何预先指定的小数还小。因此,从(6)式,可以通过选择 p 充分大而使 n 个数 ε_v 任意小。

由此立即可知,可以使 n 项之和 $a_1\varepsilon_1+\cdots+a_n\varepsilon_n$ 任意小。事实上,我们有

$$|a_1\varepsilon_1+\cdots+a_n\varepsilon_n|\leqslant|a_1\varepsilon_1|+\cdots+|a_n\varepsilon_n|。$$

而根据(6)式又有

$$\leqslant(|a_1|\cdot 1g_1+|a_2|\cdot 2g_2+\cdots+|a_n|\cdot ng_n)\cdot\frac{G^{p-1}}{(p-1)!}。$$

因为括号内的值与 p 无关,由于因式 $G^{p-1}/(p-1)!$ 的原因,我们可使整个右端,从而使 $a_1\varepsilon_1+\cdots+a_n\varepsilon_n$ 变成我们所要求的那样小,特别是小于 1。

据此就证明了我们所要证的,即设等式(3)

$$(a_0M+a_1M_1+\cdots+a_nM_n)+(a_1\varepsilon_1+\cdots+a_n\varepsilon_n)=0$$

成立会导致矛盾,即一个不为零的整数加上一个纯小数不为 0。因为这样的等式不存在,e 的超越性也就被证明了。

π 的超越性的证明

我们回头来证明 π 的超越性。这个证明比前一个证明困难些,但还算是容易的,只是必须从正确的结论逆推过去,而这一点确实是全部数学发现的艺术。

林德曼所考虑的问题如下:到目前为止,业已证明,如果系数 a_v 和指数 v 是通常的整数的话,等式 $\sum\limits_{v=0}^{n}a_v\mathrm{e}^{v}=0$ 是不能成立的。当 a_v, v 是任意代数 5 时,就不能得到同样的证明吗?他证明了这一点。事实上,他的关于指数函数的最一般的定理是:如果 a_v, b_v 是代数 5,a_v 是任意的而 b_v 彼此不相同,则等式 $\sum\limits_{v=1}^{n}a_v\mathrm{e}^{b_v}=0$ 是不能成立的。于是,π 的超越性作为本定理的推论而被证明。因为,众所周知,$1+\mathrm{e}^{i\pi}=0$,如果 π 是代数 5,则 $i\pi$ 亦然,而这个等式的成立与上述林德曼定理相矛盾。

我现在只来详证林德曼定理的一个特殊情形,由此可导出 π 的

超越性。我仍然主要是按照希尔伯特在《数学年刊》第43卷上的证明，它在本质上比林德曼的证明简单，而且正是对e讨论的推广。

出发点是关系式

$$1+e^{i\pi}=0。\tag{7}$$

如果π满足任何整系数的代数方程，则$i\pi$也满足这样一个方程。设$\alpha_1, \alpha_2, \cdots, \alpha_n$是包括$i\pi$在内的该方程的所有根，则由(7)式，必然也有

$$(1+e^{\alpha_1})(1+e^{\alpha_2})\cdots(1+e^{\alpha_n})=0。$$

展开后可得

$$1+(e^{\alpha_1}+e^{\alpha_2}+\cdots+e^{\alpha_n})+(e^{\alpha_1+\alpha_2}+e^{\alpha_1+\alpha_3}+\cdots+e^{\alpha_{n-1}+\alpha_n}) \\ +\cdots+(e^{\alpha_1+\alpha_2+\cdots+\alpha_n})=0,\tag{8}$$

其中某些指数也许偶然为0。每当出现这种情况，左边和式中将出现一个正的被加数1，我们把它们和第一项的1加在一起成为一个正整数a_0，它必然不为零。其余不为零的指数，表之以$\beta_1, \beta_2, \cdots, \beta_N$，于是将(8)式写成

$$a_0+e^{\beta_1}+e^{\beta_2}+\cdots+e^{\beta_N}=0\quad(其中\ a_0>0)。\tag{9}$$

现在，$\beta_1, \cdots, \beta_N$也是整系数代数方程的根，因为从根为$\alpha_1, \cdots, \alpha_n$的方程式，可以构成仍有整系数而根为$\alpha_1+\alpha_2, \alpha_1+\alpha_3, \cdots$的方程，然后作出以$\alpha_1+\alpha_2+\alpha_3, \alpha_1+\alpha_2+\alpha_4, \cdots$为根的方程，等等。最后，$\alpha_1+\alpha_2+\cdots+\alpha_n$自身是有理数并因此而满足一个线性整系数方程。将所有这些方程乘在一起，又得到一个整系数方程，它可能具有某些零根，但其余的根就是$\beta_1, \cdots, \beta_N$。省略去对应于零根的未知数的幂，就剩下具有N个β为根的N次整系数方程，且其绝对项不为零

$$b_0+b_1z+b_2z^2+\cdots+b_Nz^N=0，其中\ b_0, b_N\neq 0。\tag{10}$$

现在可以证明林德曼定理的下述特殊情形：如果$\beta_1, \cdots, \beta_N$是

一个整系数 N 次代数方程的根，且 $a_0 \neq 0$，则形为(9)式的等式不可能存在。这个定理包含了 π 的超越性。

这个证明的步骤和证明 e 的超越性时一样。正像在那里可以用有理数来逼近幂 $e^1, e^2, \cdots, e^n$ 一样，这里要考虑(9)式中的 e 的幂的最佳可能逼近，并用旧符号写成

$$e^{\beta_1} = \frac{M_1 + \varepsilon_1}{M},\ e^{\beta_2} = \frac{M_2 + \varepsilon_2}{M},\ \cdots,\ e^{\beta_N} = \frac{M_N + \varepsilon_N}{M}, \tag{11}$$

其中分母 M 仍然是一个普通的整数，但 $M_1, \cdots, M_N$ 不像以前一样是整数，而是整的代数数。而 $\beta_1, \beta_2, \cdots, \beta_N$ 一般来说是复数，其绝对值十分小。这一个证明比前一个证明难就难在这里。但是所有 $M_1, \cdots, M_N$ 之和仍然是整数。将它们重新排列，使得等式(将(9)式乘以 M 并利用(11)式)

$$(a_0 M + M_1 + M_2 + \cdots + M_N) + (\varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_n) = 0 \tag{12}$$

的第一个被加项是一个非零整数，而第二个被加项的绝对值仍将小于 1。本质上，这就是我们前面利用过的同样矛盾。借此矛盾证明(12)式和(9)式是不可能的，从而完成了我们的证明。至于细节，我们将再次证明 $M_1 + \cdots + M_N$ 可被某质数 p 整除，而 $a_0 M$ 却不能，从而证明(12)式的第一个和式不为 0；然后选择 p 充分大，使第二个和式任意小。

(1) 我们首先关心的是用埃尔米特积分的适当的推广来确定 M。这里有一个提示：埃尔米特的因式 $(z-1)(z-2)\cdots(z-n)$ 的零点就是所假定的代数方程里 e 的指数。因此，这个因式可代之以在(9)式中所用指数(即方程(10)的解)构成的乘积

$$(z-\beta_1)(z-\beta_2)\cdots(z-\beta_N) = \frac{1}{b_N}(b_0 + b_N z + \cdots + b_N z^N)。\tag{13}$$

其后，关键之处是加上 b_N 的一个适当的幂为因子，这在以前是没有必要的，因为 $(z-1)\cdots(z-N)$ 是整系数的。于是，最后令

$$M=\int_0^{\infty}\frac{\mathrm{e}^{-z}z^{p-1}\mathrm{d}z}{(p-1)!}(b_0+b_1z+\cdots+b_Nz^N)^p b_N{}^{(N-1)p-1}。\quad(14)$$

(2) 正如以前一样，按 z 的幂展开 M 的被积函数。于是，含最低次幂即 z^{p-1} 的项为

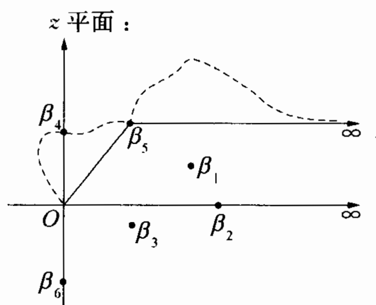
$$\int_0^{\infty}\frac{\mathrm{e}^{-z}z^{p-1}\mathrm{d}z}{(p-1)!}b_0^p\, b_N{}^{(N-1)p-1}=b_0^p\, b_N{}^{(N-1)p-1},$$

其中积分已按 Γ 函数算出。被积函数中其余各项包含 z^p 或更高次幂，因此积分后含因子 $p!/(p-1)!$ 再乘以整数，故能被 p 整除。于是，M 是一个整数且不能被 p 整除，只要 p 不是 b_0 或 b_N 的因子。但因 b_0，b_N 都不为 0，只要选择 p，使 $p>|b_0|$，$p>|b_N|$ 即可做到这一点。

因为 $a_0>0$，所以只要加上条件 $p>a_0$，即可使 a_0M 不能被 p 整除。由于有无穷多个质数，故有无限多种方法满足这些条件。

(3) 现在必须建立 M_v 与 ϵ_v。由于这里代替 v 的 β_v 可以是复数，事实上其中之一是 $\mathrm{i}\pi$，所以必须修改我们原先的计划。如果要像以前那样将积分 M 分成两部分，则必须首先确定在复平面内的积分线路。幸而，我们的被积函数是 z 的有限单值函数，除在 $z=\infty$ 处有一个本性奇点外，处处正则。我们不沿实轴从 0 到∞积分，而选任何从 0 到∞的线路，只要它最终渐近地平行于实半轴就行了。按 e^{-z} 在复平面的性质，为了使积分有意义，这样做是必要的。

现在平面上将 N 个点β_1，…，β_N 标出来。回想一下，如果首先沿直线从 0 到 β_N 中之一点积分，然后沿平行于 x 轴积分到∞(附图 3)，就可以得到同样的 M 值。沿这条线路，可将 M 分成两个各有特性的部分；沿 0 到 β_v 的直线的积分为



附图 3

ε_v，它可随 p 的增加而变得任意小；沿平行 x 轴的直线从 β_v 到 ∞ 积分，则给出代数整数 M_v

$$\varepsilon_v = e^{\beta_v}\int_0^{\beta_v} \frac{e^{-z}z^{p-1}dz}{(p-1)!}(b_0 + b_1 z + \cdots + b_N z^N)^p b_N{}^{(N-1)p-1}, \tag{14a}$$

$$(v = 1,\ 2,\ \cdots,\ N)$$

$$M_v = e^{\beta_v}\int_{\beta_v}^{\infty} \frac{e^{-z}z^{p-1}dz}{(p-1)!}(b_0 + b_1 z + \cdots + b_N z^N)^p b_N{}^{(N-1)p-1}。 \tag{14b}$$

这些假设满足(11)式。我们选择直的积分路线完全是为了方便。当然，选择从 0 到 β_v 的曲线路线，能得出同样的值，但当路线是直的时，较易估计积分。类似地，可选择从 β_v 到 ∞ 的、有水平渐近线的任意曲线来代替水平直线，但会增加不必要的麻烦。

(4) 首先讨论 ε_v 的估计，因为这一项不牵涉到任何新的内容，只要记住复积分的绝对值不大于被积函数的最大值乘以积分路线的长度，在这里是 $|\beta_v|$。这样的话，ε_v 的上极限就是 $G^{p-1}/(p-1)!$ 再乘以与 p 无关的因子。G 表示 $|\ z(b_0 + b_1 z + \cdots + b_N z^N)b_N^{N-1}\ |$ 在一区域中的最大值，而此区域包含连接 0 与 $\beta_v(v = 1,\ \cdots,\ N)$ 的整个直线段。由此，像前面一样，通过增加 p，可使 ε_v，从而使 $\varepsilon_1 + \cdots + \varepsilon_N$ 要多小就有多小，特别是小于 1。

(5) 只有在讨论 M_v 时，才有完全新的考虑，其实也只是前面推理的推广，不过现在用代数数代替过去的整数。我们将作为一个整体来考虑和

$$\sum_{v=1}^{N} M_v = \sum_{v=1}^{N} e^{\beta_v}\int_{\beta_v}^{\infty} \frac{e^{-z}z^{p-1}dz}{(p-1)!}(b_0 + b_1 z + \cdots + b_N z^N)^p \cdot b_N{}^{(N-1)p-1}$$

如果利用(7)式，用乘积 $(z-\beta_1)(z-\beta_2)\cdots(z-\beta_N)$ 来代替上面和式里每一项中 z 的多项式，并引入 $\zeta = z-\beta_v$ 这个从 0 到 ∞ 取遍实值的

新积分变量，我们得

$$\begin{cases}\displaystyle\sum_{v=1}^{N} M_v = \sum_{v=1}^{N}\int_0^{\infty}\frac{e^{-\zeta}d\zeta}{(p-1)!}(\zeta+\beta_v)^{p-1}(\zeta+\beta_v-\beta_1)^p\cdots\zeta^p\cdots \\ \qquad (\zeta+\beta_v-\beta_N)^p b_N{}^{Np-1}, \\ \text{也可写成} = \displaystyle\int_0^{\infty}\frac{e^{-\zeta}d\zeta}{(p-1)!}\zeta\Phi^p(\zeta),\end{cases} \tag{15}$$

其中我们令

$$\Phi(\zeta) = \sum_{v=1}^{N} b_N^{Np-1}(\zeta+\beta_v)^{p-1}(\zeta+\beta_v-\beta_1)^p\cdots(\zeta+\beta_v - \beta_{v-1})^p(\zeta+\beta_v-\beta_{v+1})^p\cdots(\zeta+\beta_v-\beta_N)^p。\tag{15a}$$

和式 $\Phi(\xi)$ 和它的 N 项中的每一项同样是 ζ 的多项式。在每一项中，N 个量 β_1，…，β_N 中的一个有着值得注意的地位。但如果考虑把 $\Phi(\zeta)$ 展开以后所得到的 ζ 的多项式，就看到这 N 个量同时出现在 ζ 的不同幂的系数中而未表现有何不同。换句话说，这些系数都是 β_1，…，β_N 的对称函数。通过乘法定理把这些因式乘开后可进一步推知，这些 β_1，…，β_N 的函数是具有有理整系数的有理整函数。但根据著名的代数定理，一个有理方程的所有根的有理对称函数，当其系数是有理数时，此函数本身也是一个有理数。既然 β_1，…，β_N 都是方程(10)的根，所以 $\Phi(\zeta)$ 的系数实际上都是有理数。

但是，我们所需要的是有理整数。这可由作为 $\Phi(\zeta)$ 的因子而出现的 b_N 的幂来给出。事实上，我们能把这个幂分配到其中的所有线性因式内，并写成

$$\Phi(\zeta) = \sum_{v=1}^{N}(b_N\zeta+b_N\beta_v)^{p-1}(b_N\zeta+b_N\beta_v-b_N\beta_1)^p\cdots(b_N\zeta + b_N\beta_v - b_N\beta_{v-1})^p(b_N\zeta+b_N\beta_v-b_N\beta_{v+1})^p\cdots(b_N\zeta + b_N\beta_v-b_N\beta_N)^p。\tag{15b}$$

和前面一样，把这个多项式展开后，ζ 的系数是乘积 $b_N\beta_1$, $b_N\beta_2$, …, $b_N\beta_N$ 的有理整对称函数，其系数为整数。但这个 N 个数之积是方程

$$b_0 + b_1 \frac{z}{b_N} + \cdots + b_{N-1}\left(\frac{z}{b_N}\right)^{N-1} + b_N\left(\frac{z}{b_N}\right)^N = 0$$

的根，此方程是用 z/b_N 代替(10)式中的 z 而变成的。如果用 b_N^{N-1} 乘此方程，得

$$b_0 b_N^{N-1} + b_1 b_N^{N-2} \cdot z + \cdots + b_{N-2} b_N z^{N-2} + b_{N-1} z^{N-1} + z^N = 0。\quad (16)$$

这是一个最高次系数为 1 的整系数方程，满足这样方程的数称之为整代数数。因此对上述定理可作如下改进：最高次幂系数为 1 的整系数方程，其所有根的有理整对称函数也是有理整数。你们可以在代数教材里找到这个定理，即使没有讲得这么精确，你们也可以根据证明的步骤，确信其正确。

现在，多项式 $\Phi(\zeta)$ 的系数实际上已满足此定理的假设，故皆为有理整数，记之为 A_0, A_1, …, A_{Np-1}。于是我们有

$$\sum_{v=1}^{N} M_v = \int_0^{\infty} \frac{e^{-\zeta}\zeta^p \, d\zeta}{(p-1)!}(A_0 + A_1\zeta + \cdots + A_{Np-1}\zeta^{Np-1})。$$

有了这个式子，已基本达到目的了。因为，如果利用 Γ 函数完成分子的积分，我们得到因子 $p!$, $(p+1)!$, $(p+2)!$, …，因为每一项均含有 ζ 的 p 次以上的幂。除以 $(p-1)!$ 后，仍然保持一个 p 的倍数作为因子，而其他因子是有理整数(A_0, A_1, …)。因此，$\sum_{v=1}^{N} M_v$ 必然是一个可被 p 整除的有理整数。

我们已看到，$a_0 M$ 不能被 p 整除，故

$$a_0 M + \sum_{v=1}^{N} M_v$$

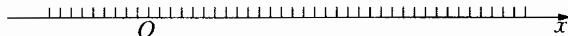
必然是一个不能被 p 整除的有理整数，特别是不为 0。因此，等式(12)

$$\left\{a_0 M+\sum_{v=1}^{N} M_v\right\}+\left\{\sum_{v=1}^{N} \varepsilon_v\right\}=0$$

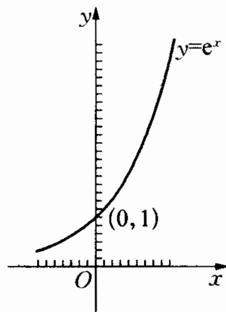
不可能存在，因为前面已证明 $\sum_{v=1}^{N}\varepsilon_v$ 的绝对值小于 1，一个不为零的整数加上 $\sum_{v=1}^{N}\varepsilon_v$ 不可能为零。这就证明了林德曼定理的前述特殊情形，亦即证明了 π 的超越性。

这里，我还想谈一下林德曼的一般定理中另一个有意义的特殊情形，即除 $\beta=0$, $b=1$ 这个平凡的特例外，方程 $e^{\beta}=b$ 的 β 和 b 不可能都是代数数。换句话说，一个代数数 β 的指数函数和一个代数数 b 的自然对数，除这种平凡的情况外都是超越数。这里就包括 e 和 π 的超越性，对前者，$\beta=1$；对后者，$b=-1$（因为 $e^{i\pi}=-1$）。对上述讨论加以确切的推广，即可证明本定理。证明可以从 $b-e^{\beta}$ 开始，而不用以前的 $1+e^{\alpha}$。这样不仅必须考虑 β 的代数方程的所有根，而且要考虑 b 的方程的所有根，才能获得一个类似于(9)式的方程，因而需要更多的符号，使证明显得难懂一些，但不需要完全新的思想。

我不打算进一步给出这些证明，但我想用几何图形指明上述关于指数函数的定理的意义。我们设想，把所有具有代数数横坐标的



点都标在 x 轴上。我们知道，有理数已是稠密的了，代数数当然更稠密。一开始有人可能会想到，代数数将穷尽所有实数，但我们的定理说明不是这种情况；在代数数之间还有无穷多个其他的数，即超越数，且我们已有无穷多个超越数的例子，即 $e^{\text{代数数}}$，log（代数数）及这些超越数的所有代数函数。如果把方程写成为 $y=e^x$，并在 xy 平面上画出曲线（附图 4），这事就更为明显了。如果在



附图 4

x 轴和 y 轴上都标出代数数，并考虑平面上 x，y 坐标都是代数数的点，这些点将“稠密地”覆盖住 xy 平面。尽管有这样稠密的分布，指数曲线 $y=e^x$ 上除去 $x=0$，$y=1$ 外没有一个代数点。所有其他满足 $y=e^x$ 的数对 x、y 中，至少有一个 x 或 y 是超越数。指数曲线的这种趋势，当然是一个最值得注意的特点。

这些定理的重大意义在于揭示出存在着大量的不仅属于有理数，而且不能用所有有理数的代数运算来表示的数，这对我们的数的连续性的概念有巨大的影响。如果说无理数的发现对毕达哥拉斯来说已是值得宰一百头牛来进行一次大祭的话，那么在作出这样的发现之后，他又该怎样去祭祀呢！

值得注意的是，尽管彻底思考一下就明白，超越数的问题是非常简单的，但一般人还是没有掌握和消化它。我曾一再遇到这种情况：在进行考试时，应考者甚至不能解释“超越”这个概念。他们常常回答我说：一个超越数不满足任何代数方程。这当然是错误的，例如 $e-x=0$ 就是一个例子。方程的系数必须是有理数这个要点被忽视了。

如果你们把上述超越性证明再彻底思考一下，你们就能从整体上掌握这些简单的步骤，并永远化为己有。你们只需把埃尔米特积分记在脑子里，其他一切都会迎刃而解。我要强调的是，在这些证明中，我们已经用到积分概念（或用几何语言来说，用到面积概念），把它看成本质上很初等的东西，而且相信这对证明的清晰性起到了很大作用。在这一方面，请对照韦伯—韦尔斯坦因第一卷里的叙述或我那本小书 *Vorträge über augewählte Fragen der Elementargeometric*（莱比锡，1895 年，F·塔格特整理）中的叙述。我想你们一定会承认，上述两本书里的证明是远远不及前述证明清晰并易于掌握，因为在上述两本书里，就像旧时的学校教科书里一样，避免使用积分符号，而用展成级数的近似计算代替了。

这些关于代数数在实数域的分布情况的讨论，自然也会把我们

引到在我的讲课过程中常常提到的第二个现代领域——集合论，我现在就对此作比较详细的讨论。

Ⅱ. 集　合　论

集合论的奠基人康托的研究，正是从考虑超越数的存在开始的[①]。他的研究使人们从一个完全新的观点来看待超越数。

如果说我对你们作的集合论简要概述有什么特点的话，那就是把处理若干具体例子放在重要地位，而不像通常那样采用十分概括的抽象说明，使人很难掌握，甚至失去信心。

1. 集合的势

为此，我提醒你们，在以往的讨论中，我们曾不得不经常涉及有不同特性的数的总和，现在我们可称之为数的集合。如果限于实数，这些集合是：

（1）正整数。

（2）有理数。

（3）代数数。

（4）所有实数。

这些集合中的每一个都含有无穷多个数。我们的第一个问题是：尽管如此，能否在一个确定意义下比较这些集合的大小或范围，即能否称某个“无穷”大于、等于或小于另一个“无穷”？康托的伟大功绩就在于通过建立准确的概念，澄清并回答了这个确实十分不确定的问题。首先我们要考虑他的势或基数的概念：当两个集合的元素能够安排得一一对应，即当两个集合能被这样联系起来，使一个集合的每一个元素，均对应于另一个集合的一个元素，反之亦然，则称

① 见 *Journal für Mathematik*，第 77 卷（1873 年）. 第 258 页。

此两集合有相等的势(等价)。如果不可能建立这样的关系,就称为不等势。如果不论用什么办法去建立对应关系,一个集合总有元素剩下来,则称此集合有较大的势。

现在我们把这个原则应用到上面的 4 个例子中。一开始,似乎会觉得正整数的势会小于有理数的势,而后者又小于代数数的势,最后小于所有实数的势,因为每一个集合都是由前一个集合加上新的元素而组成。但这个结论下得太匆忙了。因为,虽然一个有限集合的势总是大于它的一部分的势,但对无穷集合这却是不正确的。对这种差异不必惊讶,因为我们涉及的是两个完全不同范畴的情况。我们来看一个简单的例子,它表明一个无穷集合及其一部分可具有相同的势,即所有正整数和所有正偶数集合为

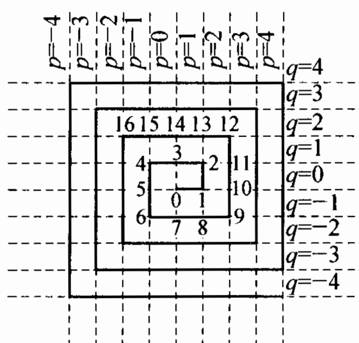
$$\begin{array}{ccccccc} 1, & 2, & 3, & 4, & 5, & 6, & \cdots, \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\ 2, & 4, & 6, & 8, & 10, & 12, & \cdots, \end{array}$$

用双箭头标出的对应显然属于前面所规定的那一类,一个集合中的每一个元素对应另一个集合的一个元素,且只对应一个。因此,按康托的定义,正整数集和它的偶数部分集有相同的势。

你们看到,上述 4 个集合的势的问题并不很容易处理。康托在 1873 年的伟大发现提供了一个简单的答案,这个答案唯其简单,更显得奇妙。答案是:正整数、有理数和代数数 3 个集有相同的势,但所有实数的集有另外一个较大的势。一个集合的元素如能与正整数集合一一对应(因而有同样的势),则称之为可数集。因此上面的定理可说成:有理数与代数数集是可数的,所有实数集是不可数的。

我们首先给出关于有理数集合的证明,你们当中有些人无疑已熟悉它了。每个有理数(包括负的),可以唯一地表达为形式 p/q,其中 p,q 是无公约数的整数,其中 q 是正的,p 可以是零或负的。为了使所有这些分数排成一个序列,在 pq 平面上标出所有具有整数

坐标(p, q)的点，使它们成为附图 5 所示的一条螺旋线上的点。于是可以数出所有的数对(p, q)，对每对只指定一个整数号码，并把所有的整数都用完(附图 5)。现在，从这个序列中删除所有不满足上述规定(p 与 q 互质，$q>0$) 的数对(p, q)，重新数剩下的(在图中用重点符号表示)，从而得到一个数列，其开头部分如下：



附图 5

$$\begin{array}{cccccccccccc} \underset{\smile}{1} & \underset{\smile}{2} & \underset{\smile}{3} & \underset{\smile}{4} & \underset{\smile}{5} & \underset{\smile}{6} & \underset{\smile}{7} & \underset{\smile}{8} & \underset{\smile}{9} & \underset{\smile}{10} & \underset{\smile}{11} & \cdots, \\ 1 & 0 & -1 & 2 & \frac{1}{2} & -\frac{1}{2} & -2 & 3 & \frac{3}{2} & \frac{2}{3} & \frac{1}{3} & \cdots, \end{array}$$

其中，一个正整数对应于一个有理数，一个有理数对应于一个正整数。这说明，有理数是可数的，这样将有理数排成可数序列，当然要完全打乱它们的大小排列，如附图 6 所示，其中横轴上的有理点，按在人为的序列中出现的顺序而排列。

有理数 p/q　−2　−3/2　−1　−1/2　0　1/2　1　3/2　2　3

正整数　7　14　3　13 6 12　2　11 5 10　1　9　4　8

附图 6

其次研究代数数。这里将限于实数，尽管把复数包含在内并不会给讨论带来更多实质的困难。每个实的代数数满足一个实整式方程

$$a_0\omega^n + a_1\omega^{n-1} + \cdots + a_{n-1}\omega + a_n = 0。$$

我们设它是不可约的，即约去左边的任何有理因子和 $a_1, \cdots, a_n$ 的任何公约数。我们也设 a_0 总是正的。于是，众所周知，每 a_n 个代数

数 ω 仅满足一个这种范式的整系数不可约方程，反之，每个这样的方程最多有 n 个实代数数的根，也许少些或甚至没有。如果能将所有这些代数方程排成一个可数序列，则显然可推知它们的根是可数的，从而推知所有的实代数数是可数的。

康托成功地解决了这个问题，他对每个方程指定一个确定的数，即其指标

$$N = n - 1 + a_0 + |a_1| + \cdots + |a_{n-1}| + |a_n|,$$

并把所有这种方程按 $N = 1, 2, 3, \cdots$ 分成可数多个相继的类。这些方程中没有一个次数为 n 或其系数的绝对值超过有限数 N，故每一类中只有有限个方程，因而只有有限个不可约方程。对给定的 N，尝试所有可能的解后，可以容易地确定出系数。事实上，对于小的 N，可以立即写出方程序列的开头部分。

现在我们来考虑，对每个指标 N，对应于不可约方程的有限个实根已确定，按大小将它们排好。首先取指数为 1 的，这样排好次序的根，然后取指数为 2 的根，等等，并按这个顺序来数。用这种方法，事实上已经指出了实代数数的集合是可数的。因为我们用这种方法数出了每个实代数数；另一方面，我们用了所有正整数。事实上，只要有充分的耐心，总可以确定出这个排列中的第 7 563 个代数数是什么，也可以确定出某一个给出的代数数在此排列中的位置，不管有多复杂。

这里，虽然在同一指标下的各代数数的自然顺序被保留了下来，但我们的“数法”再次完全打乱了代数数的自然顺序。例如，两个靠近得几乎相等的数 2/5 和 2 001/5 000，分别被指数为 7 和 7 001 分得很远，而作为 $x^2 - 5 = 0$ 的根 $\sqrt{5}$ 与 2/5 却有相同的指数 7。

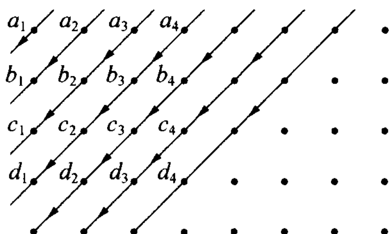
在转入最后一个例子之前，我想向你们介绍一个辅助定理，它将为我们提供另一个可数集，其证明方法对我们稍后也是有用的。如果我们有两个可数集：

$$a_1, a_2, a_3, \cdots \qquad \text{和} \qquad b_1, b_2, b_3, \cdots,$$

则由这两个集组合而成的所有 a 和所有 b 的集合也是可数的。因为可以把这个集合写成

$$a_1, b_1, a_2, b_2, a_3, b_3, \cdots,$$

并立即使它与正整数列一一对应。类似地,如果合并 3 个、4 个…或任何有限个可数集,同样可得到一个可数集。



但是我们的辅助定理是:可数无穷个可数集的并也产生一个可数集,就似乎没有这样明显了。为证明此事,我们用 $a_1, a_2, \cdots$ 排出第一个集的元素,$b_1, b_2, \cdots$ 为第二个集的元素,$c_1, c_2, \cdots$ 为第三个集的元素,如此等等。然后,只需要按上面图形所表示的方法,按照逐次的对角线选这个总集的元素,所造成的排列

1	2	3	4	5	6	7	8	9	10	11	…
a_1	a_2	b_1	a_3	b_2	c_1	a_4	b_3	c_2	d_1	a_5	…

最后会达到元素 $a, b, c, \cdots$ 中的每一个,也使它与一个确定的正整数对应,由此证明了定理。根据这个图形,可以称之为"对角线计数"过程。

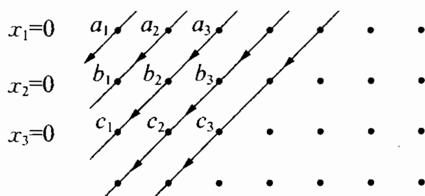
至今已知的很多不同的可数集,可能会使我们相信所有无穷集都是可数的。为了说明这是错的,我们将证明康托定理的第二部分,即所有实数的连续统是不可数的。我们用 ρ_1' 来表示连续统,因为稍后将有机会谈及多维连续统。

ρ_1 定义为所有有限实数 x 的总体,其中 x 可设想为在一条水平轴上的横坐标。我们首先指出,所有单位线段的内点 $0<x<1$ 的集合与 ρ_1 有同样的势。如果把第一个集表示在 x 轴上,把第二个集表

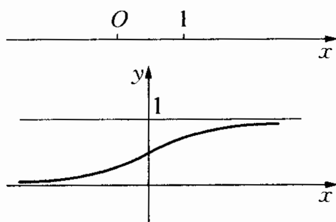
示在垂直于它的 y 轴上，则通过画在附图 7 上的那类单调上升的曲线（例如曲线 $y=-(1/\pi)\tan^{-1}x$ 的一支），就可在它们之间建立起一对一的对应。因此，可以用 0 与 1 之间的所有实数的集代表 ρ_1，今后也如此。

我将向你们介绍的关于 ρ_1 不可数的证明，是康托在 1891 年于哈勒举行的自然科学家会议上给出的。这个证明比他在 1873 年发表的那个证明更为清楚、更易推广。本质的东西是借助于所谓“对角线过程”，发现无论怎样假设所有实数的可数排列总有一个实数不可能包含在其中，是一个矛盾，因此 ρ_1 不可能是可数的。

我们把 $0<x<1$ 的所有数写成小数，并设想它们形成了一个可数数列（见第 312 页），其中 a, b, c 是在数码 0，1，…，9 中任意可能选择后再加排列的。现在必须注意我们的十进制小数不是唯一确



定的。事实上，根据我们的相等定义，我们有 $0.999\cdots=1.000\cdots$，并可以将每一个有穷小数写成无穷小数，使后者从某位起全部是 9。于是，为了得到唯一的表示，我们假设只用无穷的不尽小数，将所有有穷小数改写成以无穷个 9 结尾的小数；因而只有无穷小数出现在我们上面那个图里（附图 7）。



附图 7

现在，为了写出一个与表中每个实数都不同的实数 x'，我们把注意力集中到表的对角线上的数码 $a_1, b_2, c_3, \cdots$ 上（对角线过程一词即由此而来）。我们选择与 a_1 不同的 a_1' 作为 x' 的第一位小数，选择

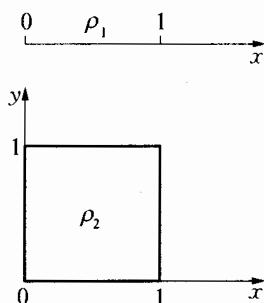
与 b_2 不同的 b_2' 作为第二位小数，选择与 c_3 不同的 c_3' 作为第三位小数，如此等等，即令

$$x' = 0.\ a_1' b_2' c_3' \cdots,$$

这里限制 a_1'，b_2'，$c_3' \cdots$ 的条件十分自由，可保证 x' 是一个真正的小数，例如不会是 $0.999\cdots = 1$，也不会在有限位数之后就终结。事实上，我们可以选取 a_1'，b_2'，$c_3' \cdots$ 使它们总是与 9 和 0 不同。这样的 x' 必然不同于 x_1，因为它们的第一位小数不同，而两个无穷小数只有在它们的每一位小数相等时才相同。类似地，由于第二位小数的原因，$x' \neq x_2$；由于第三位小数的原因，$x' \neq x_3$，等等。于是，x' 是一个不同于表内所有的数 x_1，x_2，x_3，…的一个真小数。这个矛盾的出现就证明了连续统 ρ_1 是不可数的。

本定理也先验地确保了超越数的存在，因为代数数是可数的，因此不可能穷尽由所有实数组成的不可数的连续统。但是所有以前的讨论只排出了超越数的可数无穷性，这里指出了这个集的势实际上还要大，只有这样才真正对超越数有了正确的总的认识。肯定地说，那些特殊的例子，有助于给这个图像以生命力，否则的话，就太抽象了①。

现在已经处理完了一维连续统，自然会问到二维连续统又怎么样。每个人都认为平面上的点比直线上的点要多。但康托指出，二维连续统 ρ_2 的势和一维 ρ_1 的势相同。这吸引了很多人的注意。取边长为 1 的正方形为 ρ_2，取单位线段为 ρ_1（附图 8）。我们将指出，这两个集合的点可以排成一对一的关系。这件事看起来



附图 8

① 超越数的存在性，首先是刘维尔(J. Liouville)证明的。在 1851 年法国《数学月刊》第 1 分册第 16 卷中出现的一篇文章里，他提出了一个构造此种数的初等方法。

似乎很荒谬，或许是由于我们在讨论对应时难以排除对连续性的某种心理印象。但是，我们要建立的关系要说是多么不连续都可以，或者，要说是多么杂乱无章也不为过。除去势以外，可以认为是平面和线性流形的一切特征都被打乱了，就像把正方形的所有点放入一个袋子，然后彻底搅混一样。

正方形的点集和所有纯小数对

$$x = 0.a_1a_2a_3\cdots, \qquad y = 0.b_1b_2b_3\cdots$$

一致，我们说它们都是不尽小数。我们要排除坐标(x, y)中有一为0的边界上的点，即将与原点相遇的两条边排除，保留另外两条边。很容易说明，这不会影响到势。康托证明的基本思想是，把这两个纯小数组成一个新的纯小数z，从中可以再次唯一地找到(x, y)，而当点(x, y)走遍正方形一次时，z正好取遍区间$0<z\leqslant 1$的所有值一次。如果把z想象成一个横坐标，就得到了所要求的正方形ρ_2和线段ρ_1之间的一一对应关系，而关于正方形边的约定应该包括线段中端点$z=1$。

不妨令

$$z = 0.a_1b_1a_2b_2a_3b_3\cdots,$$

从中分别选择出奇位与偶位数，即可唯一地确定出(x, y)。但由于纯小数记号含糊不清，这种做法并非无懈可击。也就是说，当(x, y)选择所有可能的不尽无穷小数对时，即走遍ρ_2'的所有点时，z可能不取遍ρ_1的全体。因为尽管z总是不尽小数，但可能存在这样的不尽小数

$$z = 0.c_1c_20c_40c_60c_8\cdots,$$

它只对应于一个有尽小数x或y，在上例中对应于

$$x = 0.c_1000\cdots, \qquad y = 0.c_2c_4c_6c_8\cdots。$$

布达佩斯的康尼格(J. König)提出的一个方法，很好地克服了这

个困难。他把 a，b，c 等不看成是单个数字而是数字复合体——可以称它们为纯小数的“分子”。一个分子是这样的小数，它由一个异于 0 的单个数字和紧靠它前面的所有的 0 组成。因此，每一个不尽无穷小数必然包含无穷个分子，因为异于零的数字必然无限次出现，反之亦然。作为一个例子，在

$$x = 0.320\,800\,700\,030\,240\,5$$

内，应取 $a_1 = 3$，$a_2 = 2$，$a_3 = 08$，$a_4 = 007$，$a_5 = 0003$，$a_6 = 02$，$a_7 = 4$，… 作为“分子”。

现在设：在上面对 x，y 和 z 之间关系的规定中，a，b，c 等作为分子而出现。于是对应于每一个数偶 (x, y)，有唯一的不尽小数 z，而 z 也随之使 x 和 y 得到确定。现在每一个 z 分成为一个 x 与一个 y，每一个 x 和每一个 y 具有无穷多个分子。因此，当 (x, y) 取遍所有无穷纯小数对时，每一个 z 正好出现一次。这就意味着单位线段和正方形已建立了一对一对应，即它们有相同的势。

当然，用类似的方法可以证明三维、四维……连续统与一维线段有相同的势。更值得注意的是，无穷维，准确地说是可数无穷维连续统 ρ_∞ 也有相同的势。这种无限维空间定义为可数无穷多个变量

$$x_1, x_2, \cdots, x_n, \cdots$$

所可能取的值的系统的总体，其中每一个变量独立地取遍所有实数。这实际上是数学上长期使用过的一个概念的新的表达形式。当我们谈到所有幂级数或所有三角级数总体时，由于系数有可数无穷多个，我们得到的实际上就是同样多个独立变量。当然，肯定地说，这些独立变量的值将受到限制，因为以计算为目的需要保证收敛性。

我们再以 ρ_∞ 里的“单位立方体”，即满足条件 $0 < x_n \leqslant 1$ 的点的全体为限，并指出它们与 ρ_1 的单位线段 $0 < z \leqslant 1$ 之间可形成一一对应关系。为了方便起见，把坐标 x_m 之一等于 0 的所有边界点和端点 $z = 0$ 除去，而保留其他边界点。和以前一样，从 ρ_∞ 里的坐标的

纯小数表示出发

$$\begin{array}{llllllll} x_1=0 & a_1 & a_2 & a_3 & \cdot & \cdot & \cdot & \cdot \\ x_2=0 & b_1 & b_2 & b_3 & \cdot & \cdot & \cdot & \cdot \\ x_3=0 & c_1 & c_2 & c_3 & \cdot & \cdot & \cdot & \cdot \\ & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array}$$

其中设所有纯小数均写成无穷形式，而且 a，b，c，…是按前面所述意义上的“十进小数分子”，即以异于 0 的数字结尾，前面有若干 0 的数字复合体。现在，必须把所有这些无穷多纯小数组成一个新的小数，而由它又可以辨认出它的成分来。或用化学语言来说，我们希望构成一种由这些分子结合成的松散的合金，且可以容易地从中分离出每一个成分。这可以通过“对角线过程”来达到。从上面的表中，按所述计划，可以得到

$$z = 0.\,a_1a_2b_1a_3b_2c_1a_4b_3c_2d_1a_5\cdots,$$

它将 ρ_∞ 中的每一个点唯一地联系着 ρ_1 的一个点。反之，用这种方法可以得出 ρ_1 中每一个点 z，因为从给定 z 的无穷纯小数，按上面给出的方法，能导出无穷个不尽纯小数 x_1，x_2，x_3，…，又用上面的方法，从中得出 z。于是，我们成功地建立起 ρ_∞ 中单位立方体和 ρ_1 中单位线段间的一一对应关系。

至今取得的结果说明，至少有两个不同的势：

(1) 可数集的势。

(2) 所有连续统 ρ_1'，ρ_2'，…，包括 ρ 的势。

自然会提出是否有更大的势这个问题。回答是：可以指出一个有更大的势的集合，这不单是抽象推理的结果，而是在数学上长期使用的概念范围之内。这个集合就是：

(3) 一个实变量 x 的所有可能的实函数 $f(x)$。

为了达到我们的目的，只要把变量限制在区间 $0<x<1$ 内就行

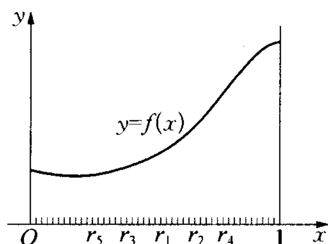
了。自然首先会想到连续函数的集合，但已有一个著名的定理。说所有连续函数具有和连续统相同的势，即属于第二类。只有把所有可以想象出来的间断函数，即在任一点 x 可以任意地取与邻近值无关的函数值的间断函数考虑在内，才可能得到更大的势。

我首先证明关于连续函数集合的定理。这就要把前面所讲过的问题重复述得圆满一些、改进一下（见第三部分，三角级数一节），使得把“任意”函数展开为三角级数的可能性更有说服力。在三角级数一节里，我曾指出：

(a) 对一个连续函数，如果知道了它在所有有理数 r 中的值 $f(r)$，它就确定了。

(b) 我们已知所有有理数可以排列成一个可数序列 r_1，r_2，r_3，…。

(c) 于是，知道可数无穷个值 $f(r_1)$，$f(r_2)$，…后，函数 $f(x)$ 就被确定了。如果要想得到一个单值连续函数，这些值当然不能任意取。这样，所有可能值 $f(r_1)$，$f(r_2)$，…的系统的集合，必然包含有一个子集，它的势与所有连续函数集合的势相同（附图 9）。



附图 9

(d) 现在，量 $f_1 = f(r_1)$，$f_2 = f(r_2)$…，可以看做是 ρ'_∞ 的坐标，因为它们是可数无穷多个连续变量的可能值的坐标。根据已证明过的定理，所有可能的值的系统的总体，具有连续统的势。

(e) 因为连续函数的集被包含在等价于连续统的一个集合之内，所以它本身必然等价于连续统的一个子集。

(f) 反之不难看到，整个连续统可以与所有连续函数集合的一部分建立起一一对应关系。为此，只需要考虑由 $f(x) = k =$ 常量确定的函数，k 为实参数。如果 k 取遍连续统 ρ_1，则 $f(x)$ 将组成一个

与 ρ_1 一一对应的集，但它只是所有连续函数集的一部分。

(g) 现在必须应用集合论中的一个重要的一般定理，即由伯恩斯坦(F. Bernstein)给出的所谓等价定理[①]：如果两个集中每一个都等价于另一集的一部分，则此两集彼此等价。这个定理是非常令人信服的，不过证明起来离题太远了。

(h) 根据(e)和(f)，连续统 ρ_1 和所有连续函数的集合满足等价定理的条件，因此它们有相同的势。定理得证。

现在转到我们的第一个定理的证明，即真正完全任意的所有可能的函数的集合有比连续统更高的势。证明立即可用康托对角线过程推出。

(a) 设定理不成立，即所有函数的集能与连续统 ρ_1 建立一一对应。设在此对应中，x 的函数 $f(x, v)$ 对应于 ρ_1 里的值 $x = v$，且当 v 取遍连续统 ρ_1 时，$f(x, v)$ 代表 x 的所有可能的函数。通过建立一个与所有这些函数 $f(x, v)$ 不同的函数 $F(x)$，说明这个假设是错误的。

(b) 为此在 $f(x, v)$ 表中建造一个"对角线函数"，即当参数 v 取值 $v = x_0$ 时，它在每一个值 $x = x_0$ 处取函数 $f(x, v)$ 在同一点的值，即 $f(x_0, x_0)$。写成 x 的函数，即 $f(x, x)$。

(c) 现在建立一个函数 $F(x)$，它在每一个 x 处的值与 $f(x, x)$ 不同：

$$\text{对所有的 } x,\ F(x) \neq f(x, x)。$$

我们有很多方法做到这一点，因为我们允许有完全间断的函数，它在任一点的值可以是完全任意的。例如，可以令

$$F(x) = f(x, x) + 1。$$

(d) 这个 $F(x)$ 确实与每一个函数 $f(x, v)$ 不同。因为如果

① 首次发表在波雷尔的 *Lecons sur la Theorie des Fonctions*，第 103 页，巴黎，1898 年。

$F(x)=f(x, v_0)$ 对某个 $v=v_0$ 成立，则当 $x=v_0$ 时等式也应成立；即应有 $F(v_0)=f(v_0, v_0)$，这与关于 $F(x)$ 的假设(c)相矛盾。

关于 $f(x, v)$ 将取遍所有函数的假设(a)被推翻，定理被证明。

把这个证明和对连续统的不可数性的类似证明比较一下，是很有趣的。那时，我们曾假设纯小数全体可以排列成一个可数表，现在我们考虑函数表 $f(x, v)$。那里选出对角线元素，相当于现在建造对角线函数 $f(x, x)$；在两种情况下，作用是相同的，即建立一个新的未包含在表内的元素，在前一情况下是一个纯小数，在后一情况下是一个函数。

你们可以很容易地想到，根据类似的考虑，可以建造出更高势的集——超过我们已讨论的 3 种势。在所有这些结果中最值得注意的是，尽管我们把这些不同的无穷集作了可以想象得到的最剧烈的处理，这种处理打乱了诸如顺序等的特殊性质，只允许最终的元素——原子——作为能以最任意的方式动荡的东西而保持其独立的存在，但在这些不同的无穷集中仍然保持着固定的区别和分类。值得指出，我们已建立起的 3 类，是数学里早已熟悉的东西——整数、连续统、函数。

我将在此结束集合论的第一部分，这里主要介绍了势的概念。我将用类似的具体的方式，更简要地向你们介绍这个理论的另外一章内容。

2. 一个集合的元素的排列

我们现在要把一个至今为止有意忽略的问题摆到面前来，即同样势的各个集合，由于其元素排列关系而产生什么样的区别。至今所接受的最一般的一一对应关系打乱了所有这些关系——只要想一下用线段代表正方形就够了。我希望强调的正是集合论的这一章的意义。集合论的目的不是想引入最一般的新概念而消除数学上早已熟悉的种种差异。相反，这个理论应通过新观点来剖析它的性质，从

而帮助我们更深刻地了解这些差异。

我们将考虑若干熟悉的例子，来搞清可能的不同排列。从可数集开始，我们曾指出本质上不同的 3 个排列的例子。这些排列是如此之不同，使其势之相同，正如我们已看到的，只是一个特殊的、绝非显然的定理的结果。这些例子是：

(1) 所有正整数的集合。

(2) 所有(正和负)整数的集合。

(3) 所有有理数和代数数的集合。

所有这些集合在其元素的排列上有一个共同性质，这个性质可以用“全序”这个名词来表达，即对两个给定的元素，总是知道哪一个在前面，或用代数语言来说，哪个大些，哪个小些。进而，如果给出了 3 个元素 a, b, c，如果 a 先于 b，b 先于 c，则 a 先于 c (如 $a<b$, $b<c$，则 $a<c$)。

现在来讲本质的不同。在(1)的情形下，有最先的元素 1，它在所有其他元素之前，但没有在所有元素之后的最后元素。在(2)的情形下，既没有最先的也没有最后的元素。(1)和(2)两个集有一个共性，即每个元素之后有一个确定的元素，每个元素之前(除去(1)中的第一个的元素)也都有一个确定的元素。与此相反，在集(3)里我们发现，在任意两元素之间总有无穷个其他元素。元素是“处处稠密”的，故在 a 与 b 之间的有理数或代数数，既无最小者，也无最大者。在这 3 个例子中的排列方法或排列类型(康托称为序型，似乎词不达意)是十分不同的，尽管势相同。这里可能会提出关于可数集的所有排列类型的问题，这正是集合论的研究者实际要回答的问题。

现在我们考虑具有连续统的势的集合。所有实数的连续统 ρ_1 是一个全序集，但在多维型的 ρ_2 和 ρ_3，…里，有非全序的例子。例如在 ρ_2 的情形下，决定两点的相互位置，需要两个关系而不是一个关系。

这里最重要的事是分析一维连续统的连续性的概念。集合理论

在澄清传统数学概念上的第一个重大成就，是认识到这里的连续性依赖于专门属于 ρ_1 的排列的简单的性质。已经发现，普通连续统的所有连续性质，其根源均在于它是具有下列两个性质的全序集：

(1) 如果我们将集合分成 A, B 两部分，使得每个元素属于两部分之一，且 A 的所有元素在 B 的前面，则或者 A 有最后元素，或者 B 有最先元素。如果回忆起无理数的戴德金定义，我们可以说，集合的每一"分割"是由集合的一个真正元素构成的。

(2) 在集合的任意两元素之间，总有无穷多个其他元素。

第二个性质是连续统和有理数可数集所共有的，但第一个性质则是两者区别的标志。在集合理论中习惯地称具有上述两个性质的全序集为连续的统，因为确实可以证明，对连续统由于连续性而成立的所有定理，对它也成立。

我提醒你们，这些连续性质也可以稍有不同地用康托的基本序列来表述。基本序列是一个集合的元素的全序可数序列 a_1, a_2, a_3, …，其每个元素或在其下一个元素之前，或在其后：

$$a_1 < a_2 < a_3 < \cdots, \qquad 或\ a_1 > a_2 > a_3 > \cdots。$$

集合中的元素 a 称为基本序列的极限元，如果(在第一种情况下)前于 a 的每一个元素最终被基本序列的元素所超过，而后于 a 的元素没有一个被基本序列的元素所超过。第二种情况也类似。如果每个基本序列在集合内均有极限点，则此集合称为闭的。反之，如果集合中每一个元素都是一个基本序列的极限元素，则此集合称为稠密的。具有连续统的势的集合的连续性，本质上就是这两个性质的联合。

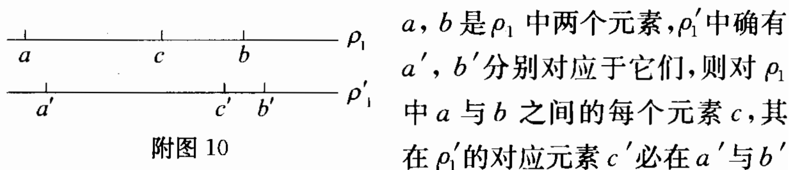
让我顺便提醒你们，讨论微积分基础时，我们谈到过另一种连续统——"维罗纳"的连续统，它是由普通连续统加上实在无穷小量而得。这个连续统由于任意两元素均有确定的顺序而成为一个全序集合，但它的排列类型与通常的 ρ_1 完全不同，甚至每个基本序列都有极限元素的定理，对它都不成立。

现在回到一个重要的问题上来，即怎样的表示能保持不同维连

续统 ρ_1，ρ_2，…之间的区别？我们确已知道，最一般的一一对应抹杀了它们的区别。这里有一个重要的定理：连续统的维数是连续一对一映射的不变量，即当 $m \neq n$ 时，不可能用一个双方单值连续的映像将 ρ_m 映成 ρ_n。有人可能轻易地接受这个定理，以为是不言而喻的道理。但你们想必记得自然直觉也几乎使我们排除可逆单值的映射 ρ_2 到 ρ_1 的可能性。这一点应使我们在接受本定理之前加以警惕。

我们只详细讨论[①]一维与二维连续统之间关系这一最简单的情形，然后指出推广到最一般情形的困难所在。我们要证明在 ρ_1 和 ρ_2 之间的双方单值连续关系是不存在的。这里每句话都是关键的。事实上，我们已经看到，不能省掉连续性。你们当中某些人无疑是熟悉"皮亚诺曲线"的例子的，这个例子也说明双方单值性不能省去。

我们需要下面的辅助定理：给出两个一维连续统 ρ_1，ρ_1'，它们被连续地相互映射，使得 ρ_1' 的每一个元素对应于 ρ_1 的唯一一个元素，且对 ρ_1 的每一个元素，ρ_1' 中最多有一个元素与之对应。于是，如果 a，b 是 ρ_1 中两个元素，ρ_1' 中确有 a'，b' 分别对应于它们，则对 ρ_1 中 a 与 b 之间的每个元素 c，其在 ρ_1' 的对应元素 c' 必在 a' 与 b' 之间(附图 10)。



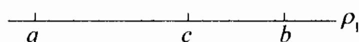
附图 10

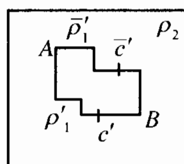
这一定理类似于下述熟悉的定理：一个在 $x = a'$，b' 取两值 a，b 的连续函数 $f(x)$，必然在 a'，b' 之间的某值 c' 处取得 a，b 之间任选的值 c。应用上述连续性的定义，可以作为这个定理的严格推广来证明这一性质。这也可以用通常定义连续函数的方法来解释连续统集合的连续映像。借助于排列的概念，可以作出这种解释。但这里不是对这些思想展开讨论的地方。

现在给出证明如下：设存在着一个将一维线段 ρ_1 映像到正方形 ρ_2 的连续的双方单值映像(附图 11)。设 ρ_1 的两元素 a，b 分别对

---

① 布劳威尔(Brouwer，L. E. J.)于 1911 年对一般情形给出了一个证明，刊于《数学年刊》第 70 卷，第 161 页。





附图 11

应于 ρ_2 中元素 A，B。现在用 ρ_2 内的两种不同路线连接 A，B，例如用画在图上的折线 ρ_1' 和 $\bar{\rho}_1'$。要做到这一点，并不需要对 ρ_2 事先假设有任何特殊性质，诸如建立坐标系等，只需要用两重序的概念。每个路线 ρ_1' 和 $\bar{\rho}_1'$ 是一个与 ρ_1 类似的全序一维连续统，因而在 ρ_1 和 ρ_2 之间的连续的双方单值关系必然使 ρ_1' 中只有一个点对应于 ρ_1' 和 $\bar{\rho}_1'$ 的每个元素，但 ρ_1 的每个元素，最多只可以对应 ρ_1' 或 $\bar{\rho}_1'$ 上的一个元素。换句话说，正因为有了上述预备定理的条件，所以在 ρ_1' 的 a 与 b 之间的每一点 c，才不仅对应 ρ_1' 中的一点 c'，也对应 $\bar{\rho}_1'$ 的一点 $\bar{c}'$。这与 ρ_1 与 ρ_2 之间的双方单值关系的假设相矛盾。定理得证。

如果试图将此考虑推广到两个任意的连续统 ρ_m，ρ_n，必须先知道能嵌入 ρ_m 中的维数为 1，2，3，…，$m-1$ 的一般连续统的构造。只要 m，$n \geqslant 2$，就不可能仅仅借助于上面最简单的情况下所用的“之间”的概念。相反，面对一些非常困难的研究，它的最初几步就已涉及平面上最一般的连续一维点集的深奥的基本几何问题，这些问题最近才稍微搞清楚了一点。其中有一个有趣的问题是，什么情况下一个点集应被称为曲线？

我们就此结束对集合论的十分专门的讨论，以便发表几点泛泛的意见。首先，想讲讲康托对点集理论在几何和分析中的地位所持的一般概念。这些概念从特殊的角度揭示了集合论的意义。算术的离散度量和几何的连续度量之间的区别，在历史上和哲学的思考上始终占有突出的地位。近来，离散量作为概念上最简单的量占了主导地位。根据这个趋势，我们把自然数、整数都看做是最简单的给定的概念，用熟悉的方法从它们推出有理数和无理数，并建立起用分析掌握几何的完整的工具，即解析几何。这种现代发展的趋势，可以称

为几何的技术化趋势。对连续性所采用的几何观点被推回到了整数的观点。我这些讲义主要坚持了这个方向。

与片面地倚重整数相反，康托想在集合论中达到“算术和几何的真正融合”(1903 年在凯塞尔的自然科学家会议上他亲自这样告诉我)。因此，整数理论，各种点的连续统理论，以及其他等等，可能在集合的一般理论中各自占有同等重要一章，形成调和的一个整体。

我再泛泛讲一下集合论与几何的关系。在集合论中，我们已经讨论过：

(1) 集合的势，在任何双方单值映像下不会改变。

(2) 考虑元素间有关顺序关系的集合的序型。这里已经能够刻画连续性的概念，不同重数的排列或多维连续统等等，使连续映像的不变量在这里找到了位置。转到几何方面来说，就是给出了从黎曼时代起就被称为位置分析(analysis situs，即拓扑学)的一个分支。这个分支是几何中最抽象的一部分，它处理在最一般的双方单值连续映像下不变的几何图形的性质。黎曼在非常一般的意义上用了“流形”这个词。康托起初也用这个词，但后来用更方便的词“集合”来代替了。

(3) 如果我们进入具体的几何学，我们将涉及度量几何与投影几何之间的区别。这里不光要知道——举例来说——直线是一维的，而平面是二维的。我们宁可去建造图像或对其进行比较。为此，我们需要用到一个固定的度量单位或至少在平面上选一条线，在空间选一个平面。在这些具体领域里，除了一般的排列性质以外，当然还需要一套专门的公理。这当然会促使全序、双序……n 重序连续集理论的进一步发展。

这里不是我详细讨论这些事情的地方，特别是因为在本书的以后几卷里必定要讲到这些。我只提一提你们以后可以参考的文献。首先我应该谈到《数学百科全书》里的 3 篇文章：恩里克斯的 *Prinzipien der Geometrie* (Ⅲ, A, B, 1)、冯·曼戈尔特(V.

Mangoldt)的 *Die Begriffe "Linie" and "Fläche"*（Ⅲ，A，B，2）（该文主要讨论公理），以及德恩—希加德（Dehn-Heegaard）的 *Analysis Situs*（Ⅲ，A，B，3）。最后那篇文章写得相当抽象。文章一开头讲拓扑学的最一般的概念以及作者本人确定的基本事实，其后一切推导皆按纯逻辑进行。这与我一贯建议的归纳式表示方法恰好相反，只有已经用归纳的方式彻底钻研过这个学科的高水平的人才能完全读懂。

至于集合论方面的文献，我应该首先提到 A · 舍恩弗列斯为 *Deutsche Mathematikervereinigung* 写的一篇文章，题为"Die Entwickelung der Lehre von den Punktmannigfaltigkeitten"①。该文第一部分发表于 *Jahresbericht der deutschen Mathematikervereinigung* 第 8 卷；第二部分最近也已发表，作为 *Jahresbericht* 附录的第 2 卷。这篇文章是整个集合论方面的真正综述，你们能从中找到许多有关的详尽资料。同时，我愿意提一提第一本系统的集合论教材：W · H · 扬与其夫人合著的《点集理论》（关于扬的夫人格雷斯 · 奇泽姆，在本书第三部分谈到球面三角学的那一节中已提到）。

在结束对集合论的讨论时，必须再次提出一个贯穿本书的问题：这方面的内容，在中学里可以采用多少？从数学教学的观点出发，我们当然必须防止过早地把这些抽象和困难的内容交给学生。为了准确地表达我对这件事的观点，我愿意摆出生物发育的基本规律，根据这个规律，个体的发展必须经历物种发展的一切阶段，不过程序缩短了。这种思想今天已变成每个人的一般文化知识的一部分。我想，数学教学和其他事情一样，都应遵循这个规律，至少是大体上遵循。教学内容要考虑到年轻人的自然能力，慢慢引导青年接触高深的事

---

① A · 舍恩弗列斯：文章的两部分发表于 1900 年及 1908 年，莱比锡。前半部分于 1913 年进行了修改，发表时题为"Entwickelung der Mengenlehre und ihrer Anwendungen"；作为此文之续，请参阅 H · 哈恩（H. Hahn）的 *Theorie der reellen Funktionen* 第 1 卷，柏林，1921 年。

物，最后接触抽象的规则。应该走人类从原来朴素的状态到高级知识阶段所走过的道路。经常阐明这个原则是必要的，因为总有这样的人，他们追随中世纪的学究方式，先教最一般的观念，还辩解说这种方法是“唯一科学的方法”。然而这个理由是完全没有事实根据的。科学的教学方法只能是促使学生去科学地思考，绝不是一开始就叫他们面对一堆枯燥的科学辞藻。

推广这种自然而真正科学的教学方法的重大障碍，是缺乏数学的历史知识，这是常常感觉到的。为了弥补这一点，我一直把历史介绍插进书中。我相信已向你们讲清，数学思想之形成是多么缓慢；这些思想最初几乎都是以近乎预言的形式，经过长期发展后，才变成严格的结晶形式，变成大家很熟悉的系统的讲法。我热诚地希望，这种知识能对你们的教学发生持久的影响。

西方数学文化理念传播译丛

丛书主编 汪 宇

Elementary Mathematics from an Advanced Standpoint

高观点下的初等数学

（第二卷） 几 何

〔德〕菲利克斯·克莱因 著

舒湘芹 陈义章 杨钦樑 译

余家荣 审



复旦大学出版社

www.fudanpress.com.cn

内 容 提 要

菲利克斯·克莱因是19世纪末20世纪初世界最有影响力的数学学派——哥廷根学派的创始人，他不仅是伟大的数学家，也是现代国际数学教育的奠基人、杰出的数学史家和数学教育家，在数学界享有崇高的声誉和巨大的影响。

本书是克莱因根据自己在哥廷根大学多年为德国中学数学教师及在校学生开设的讲座所撰写的基础数学普及读物。该书反映了他对数学的许多观点，向人们生动地展示了一流大师的遗风，出版后被译成多种文字，是一部数学教育的不朽杰作，影响至今不衰。全书共分3卷。第一卷：算术、代数、 分析；第二卷：几何；第三卷：精确数学与近似数学。

克莱因认为函数为数学的"灵魂"，应该成为中学数学的"基石"，应该把算术、代数和几何方面的内容，通过几何的形式用以函数为中心的观念综合起来；强调要用近代数学的观点来改造传统的中学数学内容，主张加强函数和微积分的教学，改革和充实代数的内容，倡导 "高观点下的初等数学"意识。在克莱因看来，一个数学教师的职责是："应使学生了解数学并不是孤立的各门学问，而是一个有机的整体"； 基础数学的教师应该站在更高的视角（高等数学）来审视、理解初等数学问题，只有观点高了，事物才能显得明了而简单；一个称职的教师应当掌握或了解数学的各种概念、方法及其发展与完善的过程以及数学教育演化的经过。他认为"有关的每一个分支，原则上应看做是数学整体的代表"，"有许多初等数学的现象只有在非初等的理论结构内才能深刻地理解"。

本书对我国从事数学学习和数学教育的广大读者具有较好的启示作用，用本书译者之一，我国数学家、数学教育家吴大任先生的话来说，"所有对数学有一定了解的人都可以从中获得教益和启发"，此书"至今读来仍然感到十分亲切。这是因为，其内容主要是基础数学，其观点蕴含着真理……"。

第二卷目录

第六部分 几何及其基础的系统讨论

第 一 版 序

在这些讲义第一卷(算术　代数　分析)的出版序中,我曾表示怀疑,讨论几何学的第二卷是否能很快出版。但是由于海林格先生的勤奋工作,本卷终于完成了。

关于这一系列讲义成书之缘由,我在第一卷序中已经讲过,没有什么特别的话要补充了。但是对于这本第二卷所采取的新形式,似乎又有必要作一番解释。

确实,这一卷的形式与第一卷太不同了。我曾下定决心,无论如何要对几何学领域作出一个综述,把我认为每一个高中数学教师应该具有的知识范围都包括进去。于是,关于几何教学的讨论就被推到每章结尾篇幅所允许的地方,前后连贯起来。

选择这种新的写作安排的一部分动机是想避免刻板的形式。但是还有更重要、更深一层的理由。几何学中没有与该学科总水平相对应的统一教材,不像在代数和分析方面那样有标准的法国教程。一个内容广的题目往往这里讲一点、那里讲一点,简直像是各个不同的研究工作者分头写的。相反,我追求的教学法目标和一般科研目标却是希望作一个比较统一的处理。

最后我希望《高观点下的初等数学》这两卷相互补充的书,会像席马克先生和我去年出版的《数学教学组织》那本讲义一样,受到教育界的欢迎。

克莱因

1908 年圣诞节于哥廷根

第 三 版 序

根据我在第一卷第三版序中讲的总计划，这一卷（第二卷）的正文及其处理未作修改，只是在细节上作了一些小的改动，并插入一些内容。原著中没有讲到的、涉及科学文献和教学法文献的两个附录，是赛法特先生和我一再讨论后编写的（中译本未收此附录——中译者）。赛法特先生又担负起了与出版有关的主要工作。海林格、弗迈尔及瓦尔特先生帮助他看了校样。弗迈尔先生编了两个索引。我十分感谢这几位先生，也对施普林格出版公司表示谢意，该公司始终如一地表现了协助出版的精神。

克莱因

1925 年 5 月于哥廷根

英 译 者 序

克莱因著《高观点下的初等数学》3 卷本第一卷英文译本出版以后，收到了读者的良好反应，所以我们翻译出版了本书，即原著第二卷。纽约大学柯朗教授在执教哥廷根大学时即已建议翻译出版克莱因的著作，他一直提供慷慨的协助，为在美国刊行第二卷铺平了道路。

前　言

先生们！我现在开始讲的课程是去年冬天课程的继续或补充。我现在的目的像那时一样，是要把你们在大学几年中学过的一切数学知识集中起来，只要对未来的教师有用就收集起来，特别是要指出它们同中学教学的关系。在去年冬天的那一学期，我已经执行了这个计划，讲了算术、代数和分析。这一学期将把注意力投到去年放在一边的几何上。在这次讲座中，课程内容是独立于上次课程的知识的。此外，我将在整体上采取有所不同的方式：先讲百科全书式的全面内容，向你们提供通盘的几何知识介绍，你们可以把已经学过的一切零星知识都纳入一个严格的系统，要用的时候就可以拿来用。作了这番通盘介绍之后，我才强调与数学教学有关的内容，而我去年冬天的出发点始终就是数学教学。

我很高兴地要提一下 1908 年复活节假日期间在哥廷根这里所举办的数理教师假期讲座。在那个讲座中，我介绍了去年冬天讲座的内容。与此有关，也由于此地中学贝伦德逊(Behrendsen)教授所作的讲话，引起了一场有趣而富有启发性的讨论，涉及了中学算术、代数、分析教学的重新组织问题，特别是谈到了把微积分引入中学的问题。参加讨论的人对这些问题表现出了极为令人欣慰的兴趣，并一般对我们使大学和中学发生紧密接触的努力表示兴趣。我希望这个讲座也会在这个方向发生一定的影响。我们以往不断地听到从中学里传来的，往往是正确的抱怨，说大学教育固然传授了许多专门内容，但对新教师以后真正会用到的许多重要的一般内容却没有讲，使新教师完全摸不清方向。但愿我这个讲座能起到一定的作用，有助

于消除这种由来已久的抱怨。

现在来讲讲这个讲座的内容。像以前的那个讲座一样，为了强调对整体内容的一般介绍，我不时需要假定你们从已经学过的一切数学知识中掌握了一些重要的定理。不错，我会始终努力作些简短的说明，以促使你们回想起学过的内容，使你们能够轻而易举地摸清文献。另一方面，我要把注意力更多地吸引到几何学科的历史发展以及伟大先驱者的成就上来，不像我在第一卷里通常做的那样。我希望，通过这类讨论，提高我常说的你们的一般数学素养，因为除了专门课程提供的详尽知识以外，还应当抓住主题内容及历史关系。

请允许我最后再泛泛地讲几句话，以免由于把这一部分几何同第一部分算术作了名义的划分而产生误解。尽管作了这种划分，但我在这里像在那种一般的讲座中一样，始终如一地最喜欢用"算术和几何的融合"这个说法来表示我的主张。我的意思是：算术这个领域不仅包含整数理论，而且包括整个代数和分析。德国中学里一般就是如此。某些人，特别是在意大利，喜欢用"融合"这个词，但仅限于指几何方面的努力。其实，无论是在大学里或在中学里，都早已形成惯例，先学平面几何，然后完全与平面几何分开来再学空间几何。因此，空间几何不幸往往遭到轻视，使我们最先具有的卓越的空间感知能力不能得到发展。与此相反，"融合派"希望同时处理平面与空间几何，使我们的思维不至于人为地受到二维的限制。这种努力也得到我的赞同，但我同时想到的是更深入、更广泛的融合。上个学期，我始终努力想使算术、代数及分析的抽象讨论生动活泼起来，利用图示和作图方法使内容更容易为个人所接受，并第一次向学生说明为什么应当对这种讨论发生兴趣。同样，现在一开始就要伴以空间观念，这种空间观念与解析公式一起当然应占首要地位，促进对几何材料的高度精确的概括。

下面我将讨论我们的问题，首先考虑一系列简单的几何基本形式，你们就能最容易地理解我讲的意思。

第四部分

最简单的几何流形

第十章
作为相对量的线段、面积与体积

看了本章的标题，你们就会了解，我准备同时考虑在直线、平面和空间里的对应变量。然而，为了分析表述，将考虑统一的原则，立即利用直角坐标系。

设有一线段，设想其位于 x 轴上。如其端点的横坐标为 x_1 和 x_2，则其长度为 $x_1 - x_2$，可以将这个差写成行列式形式

$$(1,\ 2) = x_1 - x_2 = \frac{1}{1}\begin{vmatrix} x_1 & 1 \\ x_2 & 1 \end{vmatrix}。$$

类似地，由坐标为$(x_1,\ y_1)$，$(x_2,\ y_2)$，$(x_3,\ y_3)$的 3 个点 1，2，3 形成的 xy 平面上的三角形的面积为

$$(1,\ 2,\ 3) = \frac{1}{1 \cdot 2}\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}。$$

最后，由坐标为$(x_1,\ y_1,\ z_1)$，…，$(x_4,\ y_4,\ z_4)$的 4 点形成的四面体体积的公式为

$$(1,\ 2,\ 3,\ 4) = \frac{1}{1 \cdot 2 \cdot 3}\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{vmatrix}。$$

我们通常所说的线段长度或相应情况下的面积与体积，是这几个量的绝对值，而实际上，我们的公式所提供的，远远不止这些，还给

出了依赖于所取各点顺序的一个确定的符号。在几何学中，将始终考虑这些解析式子所提供的符号，以此作为基本规则，因而必须要问包含在这些行列式内的符号的几何意义。

因此，如何选择直角坐标系，就成了一件重要的事。所以请一开始就建立一个约定，这虽然是任意的，但必须在一切情况下皆有约束力。在一维情形下，我们将认为 x 轴的正向总是指向右方。在平面上，x 轴的正向指向右方，而 y 轴的正向指向上方（图 10.1）。如果使 y 轴朝下，则必有一个本质不同的坐标系，它是前者的反射，如果不进入空间仅通过在平面内的移动，是不可能使它们相互重合的。最后，通过对平面坐标系加上一个正向指向前方的 z 轴，将得到空间坐标系（图 10.2）。选取 z 轴正向指向后方，同样会得到一个本质不同的坐标系，不可能通过在空间内的任何运动而将两个不同的坐标系重合起来。这两个坐标系分别称为“右手系”和“左手系”①。

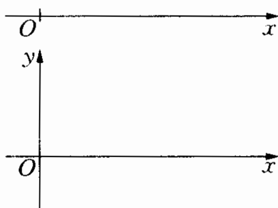


图 10.1

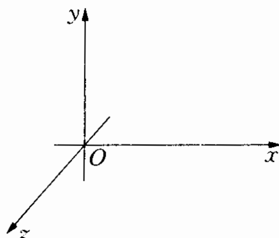


图 10.2

如果遵照这套规定，由数字标出的这些点的顺序的简单几何性质就可以用来解释我们的符号。

对于线段(1, 2)，这个性质是显然的：其长度表达式 $x_1 - x_2$ 将依赖于点 1 位于点 2 的右侧或左侧而成为正的或负的。

就三角形的情形来说，其面积公式将依赖于从点 1 经点 2 到点 3

① 这两个系统之区别，是由于相应地对应于右手和左手，开头 3 个手指的位置（见第一卷第一部分第四章第二节）。

是按逆时针或按顺时针旋转而成为正的或负的。将首先考虑一个特殊位置的三角形，计算表达其面积的行列式的值，并通过考虑连续性而过渡到一般情形，这样就证明了此结论。

我们考虑这样一个三角形，其第一个顶点为 x 轴上的单位点 ($x_1 = 1$, $y_1 = 0$)，第二个顶点为 y 轴上的单位点 ($x_1 = 0$, $y_1 = 1$)，第三个顶点为原点 ($x_1 = 0$, $y_1 = 0$)。根据关于坐标系的约定，必须沿逆时针方向走过此三角形的边界(图10.3)，而关于其面积的公式产生正值

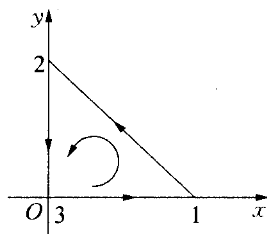


图 10.3

$$\frac{1}{2}\begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix} = +\frac{1}{2}。$$

现在，可以通过连续变形，将此三角形的各顶点与任何其他按同方向走过边界的三角形的顶点相重合，而且可以保持三角形的 3 个顶点在任何时候都不共线。在这个过程中，行列式的值连续变化，且因为只在点 1, 2, 3 共线时才化为 0，故此值必然始终为正。这就证明，任何一个沿逆时针方向走过边界的三角形的面积是正的。如果变换原三角形的两个顶点，立即可以看出，每一个沿顺时针方向走过边界的三角形的面积是负的。

可以用类似的方法讨论四面体。还是从一个特殊位置的四面体出发。依次选择第一、第二、第三顶点为 x 轴、y 轴和 z 轴上的单位点，第四顶点为原点。因此其体积为

$$\frac{1}{6}\begin{vmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{vmatrix} = +\frac{1}{6}。$$

和前面一样推出，每一个可以从这个四面体连续变形并保持 4

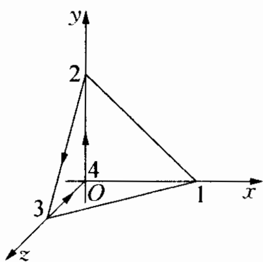


图 10.4

个顶点不共面（行列式始终不为零）而得到的四面体，其体积为正。但可以从顶点 1 看面（2，3，4）的转移方向（图 10.4），据此说明所有这一类四面体。用这种方法，我们可得如下结论：如果从顶点 1 去看顶点 2，3，4 是逆时针顺序，则公式算出的四面体（1，2，3，4）的体积是正的，反之为负的。

于是，从解析式子出发，实际上推导出了一些几何规则，对于以确定顺序给出顶点的任何线段、三角形和四面体，可指定一个确定符号。对比把长度、面积、体积考虑为绝对值的普通初等几何学，这样做有极大的好处。初等几何必须依照图形呈现的情况而区分许多情况，而现在用几个简单的一般定理就可以概括。

请从一条直线，例如 x 轴上的 3 个点分成的各线段之比这个非常简单的例子出发。为以后方便起见，我们把这 3 点记为 1，2，4（图 10.5）。我们看到，所述之比由公式 $S=\dfrac{(x_1-x_2)}{(x_1-x_4)}$ 给出，且清楚地表明

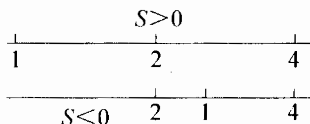


图 10.5

这个商将按点 1 在线段（2，4）之外或线段之内而成正值或负值。如果按习惯的初等表示法，只给出绝对值 $\mid S\mid=\dfrac{\mid x_1-x_2\mid}{\mid x_1-x_4\mid}$，那么总得参考图形，或是用文字说明我们所考虑的 x_2 是在线段（2，4）以内或以外，这当然要复杂得多。因此，符号的引入就把直线上各点可能的不同顺序考虑在内了，而这一点，是我在讲课过程中往往必须涉及的。

如果现在加上第四点，就可以建立起 4 点的交比，即

$$D=\frac{(x_1-x_2)}{(x_1-x_4)}:\frac{(x_3-x_2)}{(x_3-x_4)}=\frac{(x_1-x_2)(x_3-x_4)}{(x_1-x_4)(x_2-x_4)}。$$

这个表达式也有一个确定的符号，而且我们立即会看到，当点对 1 和 3 与点对 2 和 4 互相分隔时，$D<0$；相反，点 1，3 都在线段 2，4 之内或之外时，$D>0$（图 10.6、图 10.7）。因此，总是有两种本质不同的排列，产生同样的绝对值 D。如果只给出绝对值，则必须同时指明其排列情况。例如，如果仍按中学的习惯方法，用方程 $D=1$ 来确定调和点，则必须将两个点对相互分隔的要求包含在定义之内，而按这里的规定我们只要说 $D=-1$ 就够了。这种把符号考虑在内的做法，在射影几何里特别有用。正如你们所知道的，在射影几何里，交比起十分重要的作用。这里有一个熟悉的定理，即当从一个中心将在一直线上的 4 个点射影到另一直线上组成另外 4 点时，前后 4 个点的交比不变。如果现在把交比考虑为受符号影响的相对量，则此定律之逆也无例外地成立：如果位于两条直线上的两个 4 点组成相同的交比，则其中一组可以通过另一组的一次或多次射影而获得。例如，在图 10.8 里，如果用中心 P 和 P'，则点集 1，2，3，4 和 $1''$，$2''$，$3''$，$4''$ 将分别射影成 $1'$，$2'$，$3'$，$4'$。然而，如果只知道 D 的绝对值，则相应的定理不能以这种简单的形式出现，必须对各点的排列作出专门的假设。

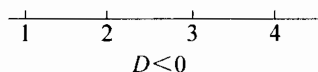


图 10.6

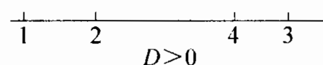


图 10.7

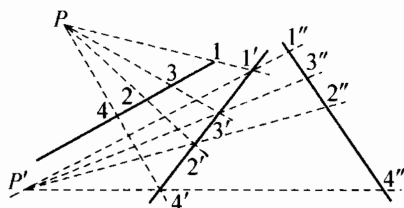


图 10.8

三角形公式的应用，是一个更富有成果的领域。首先在三角形(1, 2, 3)内部任取一点 0，并将其作为顶点之一(图 10.9)，则按初等意义下作为绝对值的 3 个小三角形面积之和等于原来三角形的面积，即可写成

$$|(1,2,3)|=|(0,2,3)|+|(0,3,1)|+|(0,1,2)|。$$

图中表明，出现在这个方程中的所有三角形顶点的顺序都是逆时针的，因此，面积(1, 2, 3)，(0, 2, 3)，(0, 3, 1)，(0, 1, 2)在我们的一般定义下都是正的，所以可以把公式写成为

$$(1,2,3)=(0,2,3)+(0,3,1)+(0,1,2)。$$

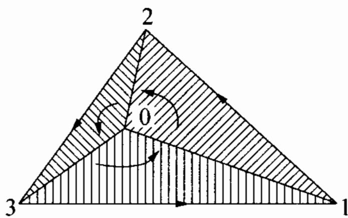


图 10.9

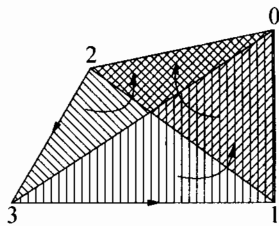


图 10.10

现在我断言，当点 O 在三角形外面，甚至当 0, 1, 2, 3 是平面上的任意 4 点时，这个公式也是成立的。以图 10.10 为例，我们看到边界(0, 2, 3)和(0, 3, 1)是按逆时针方向走过，而(0, 1, 2)则按顺时针方向走过，所以对绝对值面积，我们的公式给出

$$|(1,2,3)|=|(0,2,3)|+|(0,3,1)|-|(0,1,2)|。$$

图 10.10 表明，这个方程是正确的。

我们将借助解析定义对这个定理给出一个一般的证明。我们会看到，我们的公式是行列式论的一个著名定理。为了方便起见，不妨取 0 为原点 $x=0$, $y=0$，这显然无碍于一般性，并用合适的行列式来代替 4 个三角形的各个面积。于是，到处略去因子$\frac{1}{2}$，我们要证

明，对任意 x_1，…，y_3，下面关系式成立

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 0 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} + \begin{vmatrix} 0 & 0 & 1 \\ x_3 & y_3 & 1 \\ x_1 & y_1 & 1 \end{vmatrix} + \begin{vmatrix} 0 & 0 & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix}。$$

如果用 0 代替右边各行列式最后一列的第二行和第三行的 1，则其值不变，因为按第一行展开时，这些元素所进入的子行列式是乘以零的。如果在最后两个行列式内对行作循环变换，这对三阶（事实上对奇阶行列式）是允许的，则可以将方程式写成

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 0 & 1 \\ x_2 & y_2 & 0 \\ x_3 & y_3 & 0 \end{vmatrix} + \begin{vmatrix} x_1 & y_1 & 0 \\ 0 & 0 & 1 \\ x_3 & y_3 & 0 \end{vmatrix} + \begin{vmatrix} x_1 & y_1 & 0 \\ x_2 & y_2 & 0 \\ 0 & 0 & 1 \end{vmatrix}。$$

但这是一个恒等式，因为右端只不过是左端最后一列的子行列式，所以这里只不过是众所周知的按一列的元素展开这个行列式而已。因此，对 4 个点的所有可能情况，我们的定理一举得到了证明。

可以将此公式推广到表达任意多边形的面积。设想在测量中有这样的问题：在测量了各个角点 1，2，…，$n-1$，n 的坐标系后，确定直线多边形区域的面积（图 10.11）。一个不习惯使用符号法则的人或许会先画出多边形的草图，用对角线把它分成许多三角形，然后对某些凹进去的角作特别考虑，使欲求的面积化为各三角形面积之和或差。然而，我们可以立即给出一个不需要参看任何图形就能自然而然地得到正确结果的一般公式：设 0 为平面上任一点，例如坐标原点，则沿 1，2，…，n 绕过其边界的多边形的面积为

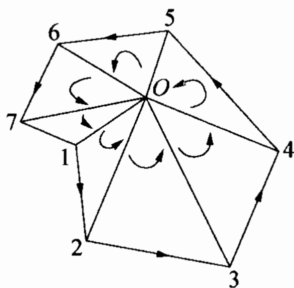


图 10.11

$$(1, 2, \cdots, n)=(0, 1, 2)+(0, 2, 3)+\cdots$$
$$+(0, n-1, n)+(0, n, 1),$$

其中每一个三角形面积所带符号由其环绕的方向确定。本公式得到的面积,依赖于多边形沿 1, 2, ⋯, n 环绕方向是逆时针或顺时针而分别取正值或负值。写出这个公式就够了,你们自己可以容易地给出证明。

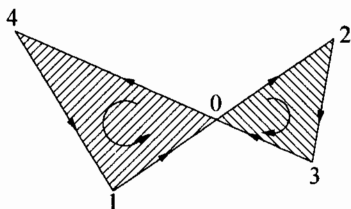


图 10.12

下一步不深入研究这个例子,而去考虑某些特别有趣的情形,这些情形事实上不可能在测量中发现,即其边界相互交叉而成为相交四边形的情形(图 10.12)。这里如果要谈确定的面积,那只能是我们的公式给出的值。让我们考虑这个值的几何意义。一开始我们已指出,它必须与点 0 的特殊位置无关。让我们把点 0 放在尽可能方便的地方,即放在相交处。于是三角形(0, 1, 2)和(0, 3, 4)化为零,而剩下:

$$(1, 2, 3, 4)=(0, 2, 3)+(0, 4, 1)。$$

第一个三角形有负面积,第二个有正面积。因此,按(1, 2, 3, 4)方向环绕时,交叉四边形的面积,等于沿逆时针方向走过的部分(0, 4, 1)的面积扣除沿顺时针方向走过的部分(0, 2, 3)的面积。

作为第二个例子,请考虑五角星形(图 10.13)。如果把点 0 放到中心部分,则在和式

$$(0, 1, 2)+(0, 2, 3)+\cdots+(0, 5, 1)$$

里所有小三角形都按正方向走过,它们的和包括图形的五角中心部分面积的两倍,5 个尖角部分面积的一倍。如果再考虑环绕此多边形的一个正向循环

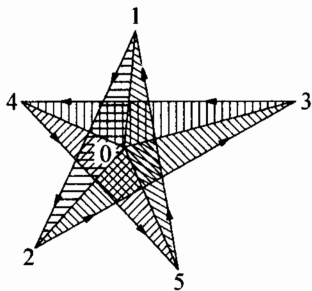


图 10.13

(1, 2, 3, 4, 5, 1),我们看到,边界的每部分都沿逆时针走过,特别是,我们环绕计算了双倍面积的多边形部分走过两次,而环绕其余部分仅一次。

从这两个例子,可以推出下面的一般规则:对边可任意相交的任何直线多边形,我们的公式所给出的总面积为由多边形的边所包围的各部分面积的代数和,其中第一部分的面积按沿(1, 2, 3, …, n, 1)环绕一次时经过的次数而计算其倍数,正负号则按逆时针或顺时针走过此部分面积而定。你们不难证明这个定理的正确性。

现在从多边形转到具有曲线边界图形的面积。我们将考虑任何封闭曲线,包括它可以与自己相交任意次数。我们指定一个沿此曲线的确定方向,并考虑由此曲线所包围的面积。如果通过不断增加边数并缩短边长使折线逼近曲线(图 10.14),用我们刚才说明的方法求出这些多边形的面积并过渡到极限,自然而然就找到其面积。如果

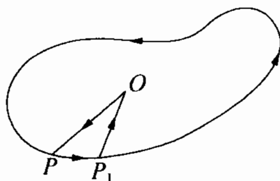


图 10.14

$$P(x,\ y)\ \text{和}\ P_1(x+\mathrm{d}x,\ y+\mathrm{d}y)$$

是这样一个逼近多边形的两相邻顶点,则其面积为基本三角形(OPP_1)面积,即

$$\frac{1}{2}\begin{vmatrix} 0 & 0 & 1 \\ x & y & 1 \\ x+\mathrm{d}x & y+\mathrm{d}y & 1 \end{vmatrix} = \frac{1}{2}(x\mathrm{d}y - y\mathrm{d}x)$$

的和。过渡到极限,此和化为沿此曲线的线积分

$$\frac{1}{2}\int(x\mathrm{d}y - y\mathrm{d}x)。$$

因此,此积分也就确定了由曲线所包围图形的面积。如果希望解释此定义的几何意义,可将刚才对多边形所得到的结果应用到新的情

况：当给定的曲线按指定的方向走过一次时，每一个被曲线包围的部分区域的面积按其沿逆时针方向被包围的次数计算正倍数，而按顺时针方向被包围的次数计算负倍数。对如图 10.14 所示的简单曲线，积分相应地给出由曲线包围区域的取正号的面积。对图 10.15 的情况，外面部分按正号计算一次，里面部分则计算两次。在图 10.16中，左部是负的而右部是正的，故计算结果为负。在图 10.17 中，一部分根本没有计算，因为它按正向和负向各被包围一次。当然，按这里的意义，可能出现这样的曲线，其包围的面积为零。如果在图10.16中使曲线相对其交点对称，就得到这样的曲线。只要记得我们是按一个方便的假定来确定面积的，那么这种情况就不足为奇了。

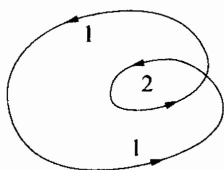


图 10.15

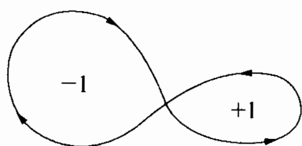


图 10.16

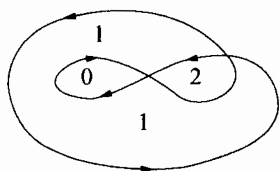


图 10.17

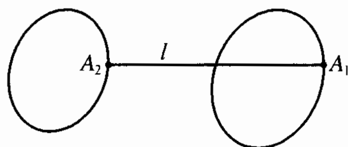


图 10.18

现在向你们指出，使用安姆斯拉(Amsler)极性求积仪，可见这些定义是多么恰当。这个极精巧而有用的工具是 1854 年由德国技工雅哥比·安姆斯拉(Jacob Amslen)制造的。用这个工具测定的面积，正好是上面讨论中所说的面积。让我们首先考虑这个仪器的理论基础。

设想一条长为 l 的棒 A_1A_2（图 10. 18）在平面上移动，使得 A_1A_2 画出分离的封闭曲线，而棒本身返回到原来的位置。我们希望求出棒所扫过的面积，把这个面积的若干部分按扫过的方向不同而以正或负计算。为此，代替棒的连续移动我们使棒从一个位置 12 到一个邻近位置 $1'2'$ 经任意小的逐次基本移动，以及在积分中所用的极限过程。棒所扫过的实际面积，为这些移动所形成的“基本四边形”$(1, 1', 2', 2)$ 面积的和的极限。很容易看到，对应于 $1, 1', 2', 2$ 的巡回方向，为每个基本四边形面积给出符号，也就把棒的移动方向适当地考虑进去了。现在，可以把棒 A_1A_2 的每个基本移动分解为 3 步（图 10. 19）：

(1) 沿棒所在方向平移 $\mathrm{d}s$。

(2) 沿棒的垂直方向平移 $\mathrm{d}p$。

(3) 绕端点 A_2 旋转角度 $\mathrm{d}\varphi$。

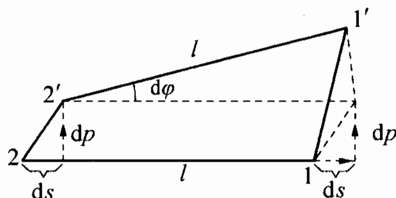


图 10. 19

按此方法，要分别扫过面积 $0 \cdot \mathrm{d}s$，$l \cdot \mathrm{d}p$，$\frac{l^2}{2}\mathrm{d}\varphi$。可以用这 3 个面积的和来代替基本四边形的面积，因为这样造成的误差是一个高阶无穷小，过渡到极限（实际上是一个简单的积分过程）则将消失。关键的是注意这个和。如果沿逆时针方向测 $\mathrm{d}\varphi$ 为正，且绕 φ 增加的平移 $\mathrm{d}\varphi$ 为正，则

$$l \cdot \mathrm{d}p + \frac{l^2}{2}\mathrm{d}\varphi$$

与四边形 $(1, 1', 2', 2)$ 的面积符号一致。

沿移动的路径积分，得到 A_1A_2 扫过的面积为

$$J = l\int \mathrm{d}p + \frac{l^2}{2}\int \mathrm{d}\varphi。$$

积分 $\int \mathrm{d}\varphi$ 表示棒对于其原来位置所转过的总角度。因为棒回到原来位置，除非做了一个完全的旋转，否则 $\int \mathrm{d}\varphi = 0$，于是面积为

$$J = l\int dp。\tag{1}$$

但是,通过适当选择 A_1 和 A_2 的路径,棒在回到原来位置之前,是有可能作一次或多次完全旋转的,于是,$\int d\varphi$ 是 2π 的倍数。对(1)式右端,相应于正向和反向的一个完全旋转,得加上 $+\pi l^2$ 或 $-\pi l^2$。为简单起见,我们把这个稍微复杂一点的情况放在一边。

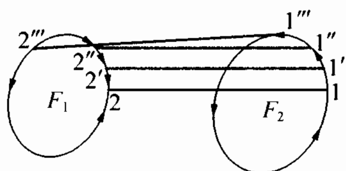


图 10.20

现在可以用稍微不同的方法来确定面积 J(图 10.20)。在逐次的基本移动中,设棒依次取得位置 12, 1′2′, 1″2″, …,则 J 是这些基本四边形面积之和:

$$J = (1,\ 1',\ 2',\ 2) + (1',\ 1'',\ 2'',\ 2') + (1'',\ 1''',\ 2''',\ 2'') + \cdots。$$

说得精确一点,是代表这个和式极限的积分,而这里所指的每个四边形环绕的方向和前面所说的一样。利用前边的多边形的公式,其中 0 为任意选定的坐标原点,我们有

$$\begin{aligned} J = &(0,\ 1,\ 1') + (0,\ 1',\ 2') + (0,\ 2',\ 2) + (0,\ 2,\ 1) \\ &+ (0,\ 1',\ 1'') + (0,\ 1'',\ 2'') + (0,\ 2'',\ 2') + (0,\ 2',\ 1') \\ &+ (0,\ 1'',\ 1''') + (0,\ 1''',\ 2''') + (0,\ 2''',\ 2'') + (0,\ 2'',\ 1'') \\ &+ \cdots, \end{aligned}$$

其中每一行的第二个三角形与下一行的第四个三角形一样,但环绕的方向相反

$$(0,\ 1',\ 2') = -(0,\ 2',\ 1'),\ (0,\ 1'',\ 2'') = -(0,\ 2'',\ 1''),\ \cdots$$

所以这些被加项全部抵消。又因为这些基本四边形系列是闭合的,被加项(0, 1, 2)将会出现在最后一行而与第一行的(0, 2, 1)相抵消,于是每一行中只留下第一、第三两个三角形。但按前面讲的,第

一个三角形面积之和为多边形(1, 1′, 1″, …)的面积,过渡到极限,即为棒的端点 A_1 所画曲线包围的面积 F_1。类似地,如果逐项改变符号,第三个三角形面积之和为多边形(2, 2′2″, …)的面积,过渡到极限,即为 A_2 所画曲线包围的面积 F_2。于是有

$$J = F_1 - F_2。\tag{2}$$

显然,只要注意考虑我们的符号规则来确定 F_1 和 F_2,两曲线可任意地相交。

上述求积仪的几何原理,就包含在(1)式和(2)式里。即:如果使 A_2 沿一个已知面积 F_2 的曲线移动,而在 A_1 处的轨迹点使其沿 F_1 的边界移动,并有一个仪器使我们能测量 $\int \mathrm{d}p$,则可立即求得值

$$F_1 = F_2 + l\int \mathrm{d}p。\tag{2′}$$

安姆斯拉所造的仪器就是这样的一个装置。他在作为轴的棒 A_1A_2 上固定一个轮子,随棒的移动而在纸上滚动。这是他的机械发明的第二部分。设轮子半径为 ρ,距离 A_2 为 λ(图 10.21)。轮子随棒移动而滚动过的角 ψ 是各个基本移动中滚过角 $\mathrm{d}\psi$ 的和,每

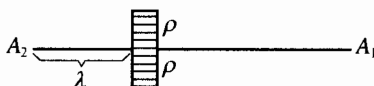


图 10.21

个 $\mathrm{d}\psi$ 可以由 3 个滚动角 $\mathrm{d}\psi_1$, $\mathrm{d}\psi_2$, $\mathrm{d}\psi_3$ 组成,每一个滚动角各自对应于前面所分解成的 3 个简单移动。在第一步移动中,轮子不转,故 $\mathrm{d}\psi_1 = 0$;在第二步,棒 A_1A_2 沿其垂直方向移动 $\mathrm{d}p$ 时,轮子在纸上转过距离 $\mathrm{d}p = \rho\mathrm{d}\psi_2$,故 $\mathrm{d}\psi_2 = \dfrac{\mathrm{d}p}{\rho}$;第三步,棒环绕 A_2 转过角 $\mathrm{d}\phi$,轮缘转过距离 $\lambda\mathrm{d}\phi = \rho\mathrm{d}\psi_3$,故$\mathrm{d}\psi_3 = \dfrac{\lambda}{\rho}\mathrm{d}\phi$。于是有

$$\mathrm{d}\psi = \frac{1}{\rho}\mathrm{d}p + \frac{\lambda}{\rho}\mathrm{d}\phi。$$

如果 A_1A_2 回到原来位置的过程中不作完全旋转,则沿整个移动路径积分之,有:$\int d\phi = 0$。于是,安姆斯拉轮子的全部转动角将是

$$\psi = \frac{1}{\rho}\int dp。\tag{3}$$

但如果棒作了一次或多次旋转,则在右侧将出现 $2\pi\frac{\lambda}{\rho}$ 的某个倍数。但我们仍不考虑此种情况。

联合(2′)式与(3)式,最终得到

$$F_1 - F_2 = l \cdot \rho \cdot \psi,$$

即由轮子转过的角 ψ 可测量出棒的两端所迂回包围面积之差。

在制作此仪器时,使 F_2 为零是可取的。安姆斯拉用一个聪明的方法把 A_2 附在一个绕固定点 M 转动的臂上(图 10.22)。于是 A_2 只能在臂上沿弧作来回移动而不会包围任何圆面积,如果我们不考虑 A_2 绕 M 作一次或多次完全的转动这个复杂的可能性的话。由于有这个“极点”M,整个仪器常常被称为极性求积仪。这个仪器的实际操作是在 A_1 处装一个标记用铅笔,使其沿希望测量的面积的图形边界移动一周,然后读出轮子转过的角 ψ,从而求得被包围的面积为 $F_1 = l \cdot \rho\psi$。仪器的常数 $l \cdot \rho$ 可以通过测量已知面积例如单位正方形而求得。

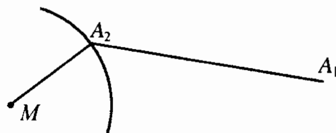


图 10.22

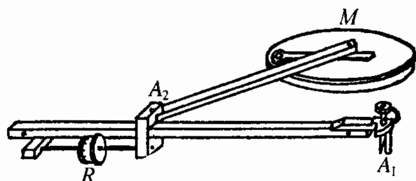


图 10.23

这里向你们介绍一个极性求积仪的图形(图 10.23)。当然,你们必须亲自去看一下,如果想完全了解它,最好去用一用。自然,如果要使仪器可靠地运行,制作方式必须比理论讨论中所说的要稍复

杂。在这一方面，我还想多说几句。点 M 用一个重物带着，并用一个杆与 A_2 连接起来。我们谈到过的，在理论上有重要意义的棒 A_1A_2，并不是你们在仪器上所看到的第二条金属杆，而是作为轮子轴的理想延长线，它与杆平行，并通过移动的铅笔点 A_1。这个尖点，用一个平行的钝的木钉带着，使点 A_1 不致把纸撕坏。轮子上带有一个可读到好几位数的游标尺和一个记录圈数的计数器。

我不再详谈了，只想提出一个一般的警告：为了说明某一理论而考虑这类仪器时，不可忽略它们的实用情况。纯数学家常常容易忽略这一面，正好像只埋头于制作仪器而对理论没有兴趣的机械师走另一极端一样，都是片面的。应用数学应该把两者结合起来。特别是应该考虑到，理论原则在仪器上是绝不能严格实现的，如仪器的连接处总有一些松动，轮子转动时总会有些滑动，画图纸不会是一个均匀平面，且铅笔不可能绝对沿曲线移动。误差多大是严重的，轮子读数的多少是可靠的等等，当然是实用中的重要问题。研究这些问题是应用数学的领域。

谈到上面讲的图形，我要说一说这些讲座同以前两个题目相似的课程的相对地位。那两个课程同样是油印的，名为：《微积分在几何上的应用：原理的修正》[①]（SS，1901 年，C · H · 缪勒（C. H. Müller），整理），以及《高等几何导论》[②]（WS，1892—1893 年及 SS，1893 年，席林（Schilling）夫人整理）。在第一个课程中，突出了刚刚提到的抽象几何和实用几何之间的区别，实际上已对安姆斯拉极性求积仪误差之根源进行了集体研讨。但在第二个课程中，我比较彻底地发展了抽象几何的理论，以满足想要以今天的研究精神在这个领域进行独立探讨的专家的需要。在现在这个课程中，我想做第三件事：阐明初等几何理论，即未来的每一名教师无疑都应了解的那些东西，特别是对于物理及力学应用具有基本意义的内容。属于前面提到的前两

---

① 新印本，莱比锡，1907 年（将于此书本版第 3 卷出版后不久出版）。

② 分两部分。新印本，莱比锡，1907 年（原书已售缺，关于新版的计划，见第 1 卷序言）。

个领域的内容,在这个课程中只能偶尔提到。

现在回到关于面积与体积的一般考虑上来。我先提一段历史。我想提到第一个在几何学上坚持使用符号原则的人——莱比锡的伟大数学家莫比乌斯。1827 年,他写了一本书,书名为《重心的计算》①。这是新几何学的一本奠基作,因下述与重心有关的考虑而得名。设在平面 3 个固定点 O_1, O_2, O_3 上放有质量 m_1, m_2, m_3,它们如同电荷一样,可以是正,也可以是负。于是重心 P 被唯一确定,我们可以通过改变 m_1, m_2, m_3 而使它们取得平面上的任何位置(图 10.24)。现在,3 个质量被当作 P 的坐标,使得 P 只依赖于这些量之比。这就是把现在所谓三线坐标引入几何的第一个例子。在那本书里,莫比乌斯把符号原则应用于确定三角形的面积和四面体的体积,并给出了我们所述的定义。我还要说及,1858 年,莫比乌斯已是一个老年人了,他又推广了这些结果,成为一个影响深远的新发现,不过这个发现直到 1865 年才以"论多面体体积的确定"②为题第一次发表。在这篇论文里,他证明了存在着不能用任何方法指定其体积的多面体。而如我们早先看到的,以任何复杂的方法相交的平面多边形,其面积皆可确定。

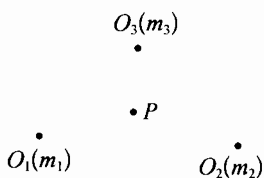


图 10.24

请从前面建立的四面体体积公式

$$(1,2,3,4)=\frac{1}{6}\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{vmatrix}$$

① 莱比锡,1827 年,或参阅其《著作选集》,第 1 卷,第 633 页,莱比锡,1885 年。

② *Berichte über die Verhand lungen der Königlich Sächsischen Gesell schaft der Wissenschaften* (Mathematischphysikalische Klasse),第 17 卷(1865 年),第 31 页,或其《著作选集》,第 2 卷,第 473 页,莱比锡,1886 年。

出发。如果按最后一列代数余子式展开这个行列式，其结果和早先在三角形的情形下所看到的一样，使一个四面体分解为 4 个四面体，它们以原点为公共顶点，而以原四面体的各个面为各自的底。按行列式理论里的符号规则，如果取循环次序为 1，2，3，4，则得公式

$$(1, 2, 3, 4) = (0, 2, 3, 4) - (0, 3, 4, 1) + (0, 4, 1, 2) - (0, 1, 2, 3)。$$

在三角形中只出现加号而这里却出现减号的理由，是在循环交换的情况下，偶阶行列式改变符号而奇阶行列式不变。当然通过适当交换行数，可以避免减号，但这样就必须放弃循环次序。例如可写成

$$(1, 2, 3, 4) = (0, 2, 3, 4) + (0, 4, 3, 1) + (0, 4, 1, 2) + (0, 2, 1, 3)。$$

为了理解这里出现的规律，设想四面体的各面由纸做成，并折平放在平面(2，3，4)上，顶点 1 则占 3 个不同位置(图 10.25)。于是，在最后一个公式中，3 个面中每一个面的 3 个顶点，如图所示，对应于逆时针方向绕所有三角形环行。

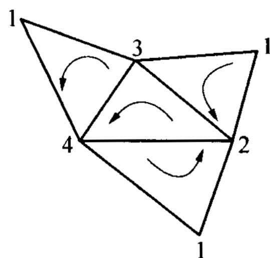


图 10.25

当然，不把各个面折平，对这个图形也能得到同样的结果。6 条棱边中每一条棱边对应于两个面，显然，当沿所有三角形按所示的顺序环行时，每条边沿相反方向各经历一次。按莫比乌斯所谓的这种棱边规律，显然只要对一个面三角形任意选定一个环行方向，则所有面三角形均有确定的环行方向。现在公式为：一个四面体(1，2，3，4)可以看做是 4 个具有公共顶点的四面体之和，只要对一个三角形(2，3，4)选定环行方向之后，按莫比乌斯棱边规律来选择其他面的环行方向。

和前面把多边形分解为三角形并推广三角形公式，以确定任意三角形的面积公式一样，现在，将设法通过刚才得到的结果以确定一个任意多面体的体积。然而，在现在的情况下，不仅应允许多面体各单个多边形面的边与其他边相交，还必须允许各个面以任意方式相截。现在选择一个任意辅助点 O。作为第一步，确定以 O 为顶点、以多面体各多边形面之一为底面的棱锥的体积。

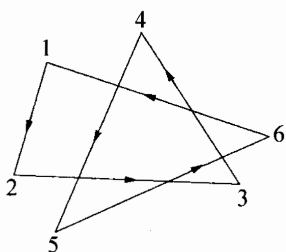


图 10.26

为此，首先对底选好方向（图 10.26），假设取多面体的面为(1，2，3，4，5，6)。按前面方法，这个多边形有确定的面积，令棱锥的体积像在初等几何中那样等于$\frac{1}{3}$的底乘高，并从点 O 看(1，2，3，4，5，6)的环行方向是逆时针方向或顺时针方向而加上正号或负号。很容易看到，这个定义把原先关于四面体体积的定义作为特例而包含在内。而且，如果用组成多边形的三角形代替该多边形，并按其环行方向使各三角形面积之和为多边形面积。然后定义棱锥的体积为以这些三角形为底的四面体体积之和，则又可以从四面体体积这个特例推出本定义。

为了在一般的意义上把多面体表示为各部分四面体之和，必须对其每一个面指定一个确定的环行方向，且必须遵循前面所说的棱边规律来做这种选择：对一个面任选一个环行方向，然后在保持两个相邻面的公共边沿相反方向各环行一次的情况下指定环行方向。如果整个多面体的面都可以这样环行而无矛盾的话，则多面体的体积可确定为：以任一点 O 为公共顶点，以按指定方向环行的多面体的面作为底的各多棱锥体体积之和。不难看到，这个定义是唯一的，且与点 O 的位置无关。

但十分值得注意的是，这个棱边规律并非适用于每一个封闭多面体的面而无矛盾，即存在着无法指定环行方向的多面体，因而不能

指定其体积。这就是莫比乌斯于 1865 年所发表的论文中的伟大发现。他在那篇论文里讨论了后来被称为莫比乌斯带的面。这个面用一张细长的矩形纸 $A_1B_1A_2B_2$(图 10.27)组成，在扭转半圈后把两端合并在一起，使 A_1 与 A_2，B_1 与 B_2 重合。显然，这样就将纸条的前后面连通起来，使这个面只有一侧。可以把它描述如下：想油漆这条带子的油漆工会发现，需用的油漆，为他从带子长度预计所要用油漆的两倍，因为油漆完带子的长度后，他会发现到了出发点的反侧，只好再环绕一次，以到达出发点。



图 10.27

我们可以把这条纸带分成几个三角形并沿其边折起，建立一个各平面部分具有同样性质的多面体表面(不封闭)。对这样得到的三角形带，棱边规律不适用。至少需要 5 个三角形，并按图 10.28 那样

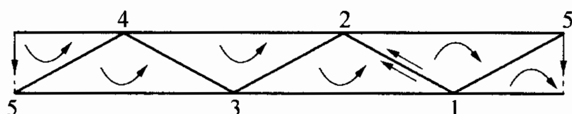


图 10.28

安排，其中右和左的两半三角形在折的过程中形成一个三角形(4, 5, 1)。这里如果选(1, 2, 3)为正的绕行方向，并按棱边规律向左延续，则顺序得到方向(3, 2, 4)，(3, 4, 5)，(5, 4, 1)，(5, 1, 2)，于是边 12 最终按与(1, 2, 3)的相同方向环行，这与棱边规律矛盾。从上面来看，这个折起来的条带，呈现为一个以 5 条边 13, 35, 52, 24, 41 为对角线的五角形，如图 10.29 所示。莫比乌斯用这个三角形组成的带子构成一个封闭的多面体。他的方法是：在五边形中间的上方，适当地选一任意空间点 O，以此为公共顶点，用带有这个公共顶点的几个三角形，把多面体的自由棱边(5 条对角线)连接起来。换句话说，这是一个具有相截表面的五侧棱锥体。当然，对这个具有 10 个三角形面的封闭多面体，也不可能应用棱边规律，所以

无法谈及它的体积①。

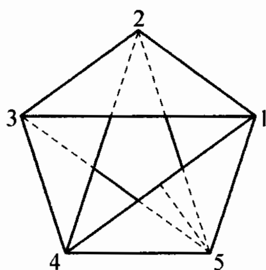


图 10.29

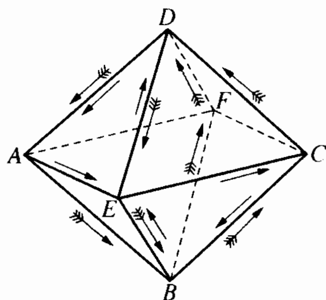


图 10.30

从八面体 $ABCDEF$(图 10.30)出发,用下面的方法可以很容易地得到另一个构造简单而封闭的单侧多面体。从八面体内选 4 个不相邻的面,相互有一个公共顶点而无公共边,例如 AED, EBC, CFD, ABF 和 3 个对角平面 $ABCD$, $EBFD$, $AECF$。这样构成的七面体②与八面体有相同的棱边,因为在后者的每条棱边中,七面体的两个相邻面相遇(即八面体的一个面和一个对角面)。八面体的各对角线不能成为此七面体的棱边,因为各对角平面不相邻。对角线 AC、BD、EF 是七面体的面相截的线。可用棱边规律证明这个七面体的单侧性。如果我们依次取出面 AED, $EDFB$, ECB, $ABCD$,对第一个面指定环行方向,并按棱边规律确定其他面的方向,最后看到 AD 按相同方向走过两次。

这里就结束以数作为容量测度的讨论,而转到其他初等几何量的处理,正好像我们在前面一直以莫比乌斯的思想为指导一样,后面

① 参阅我的文章:“Ueber Selbstspannung ebener Diagramme”,《数学年刊》,第 67 卷,第 438 页。请比较那篇文章中提到的这种单侧多面体在作图中的应用。也可参阅我的《数学著作集》,第 2 卷,第 629 页,柏林,1922 年。

② C·莱因哈特(C. Reinhardt)第一个在文献中提到,见“Zu Mobius' Polyederthorie, Verbandlungen der Königlich Söchsischen Gesellschaft der Wissenschaften”(Mathe matisch-physikalische klasse),第 37 卷,1885 年。

我们将追随什切青的伟大几何学家海曼·格拉斯曼(Hermann Grassmann)[1]在1844年发表的*Lineale Ausdehnungslehre*(《延拓理论》)中所建立的思想。这本书像莫比乌斯的那本书一样具有丰富的思想,但与莫的写作风格不同,非常晦涩,以至几十年未被人注意,也没有被人读懂。只是在其他书和文章中出现了一系列类似的思想之后,才认识到这些思想出自格拉斯曼的书,不过为时已晚。如果你想领略一下这种抽象的笔法,你只要看一下这本书里的某几章的标题,如:"纯数学之概念之导出"、"延拓理论之推导"、"延拓理论之叙述"、"表示之形式"、"一般形式理论之概述"。你只有费劲地钻通了这些内容之后才接触到所述内容的纯抽象的表示,不过仍然很难读懂。直到1862年该书出版了后期的修订本[2],格拉斯曼才用了一种比较容易接受的表示法,即坐标表示法。此外,格拉斯曼选了一个词——Ausdehnungslehre(延拓论),用以暗示他的研究可应用于任意维空间,而几何学对他而言只不过是这个完全抽象的新学科在普通三维空间中的应用。但是他造的这个新词并没有生根,人们现今简称为"n维几何学"。

现在让我们利用我们所熟悉的解析坐标以了解格拉斯曼的概念。我们首先以平面几何为限,用"格拉斯曼原理"作为下一章的标题。

---

① 海曼·格拉斯曼,《延拓理论》出版于1844年莱比锡。并可参阅其*Gesammelte mathematische und physikalische Werke*,第1卷,莱比锡,1894年,第二版出版于1898年莱比锡。

② 柏林,1862年。见其著作集第1卷第二部分,莱比锡,1896年。

第十一章
平面上的格拉斯曼行列式原理

让我们回忆一下前一章的基本解释。在前一章里，用 3 个点的坐标建立起行列式

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix},$$

并把它解释为三角形面积的两倍，即平行四边形的面积。现在请再考虑分别由两点和一点构成的形

$$\begin{pmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{pmatrix} \text{或} (x_1 \quad y_1 \quad 1),$$

并称之为矩阵。每一个这样的矩阵，代表着从中分别删去一列或两列可得到的行列式的总体。从第一个矩阵，删去第一列，然后删去第二列，可得两个二阶行列式

$$Y = y_1 - y_2, \qquad X = x_1 - x_2,$$

删去第三列，则得到行列式 $N = x_1 y_2 - x_2 y_1$。这样选择符号，是为使它也适用于空间几何。我们必然会问：从 3 个行列式 X，Y，N 所确定的几何图形是什么样的？我们将把这个图形看做是和三角形面积同样合理的一个新的初等几何量。从第二个一行矩阵中，除数 1 以外，得到一行行列式，即坐标(x_1, y_1)本身。由此确定以这些坐标作为最简单的基本量的点，这些坐标不需要进一步研究。

如果对格拉斯曼原理作出一般的说明，就可以理解：在平面上

或在空间中，我们考虑所有行数少于列数的矩阵，其中各行由一个点的坐标和 1 组成。我们问，从这些矩阵中删除适当的列而得到的行列式所确定的几何量是什么？

在这个由点任意地建立起来的原则里，我们最终会认识到，将自然而然地把许多思想推广开来，波及整个几何系统，逐渐地表现出它是认识一大堆初等几何量的有用的指导原则。

还是回到具体问题上：如果知道行列式 X，Y 和 N，在两点 1 和 2 的图形(图 11.1)中又包含着什么？显然，在两个点的位置上，仍然存在一个自由度，因为要用 4 个量才能固定它们。我要断言，当且仅当 1 和 2 分别是一条具有确定长度和方向并在一条确定的直线上自由移动的线段的终点和起点时，我们会得到同样的 3 个值 X，Y，N。从这里起，我们想象箭头从起点 2 指向终点 1。

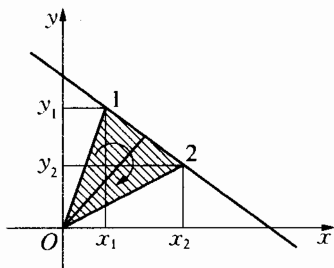


图 11.1

从方程

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0$$

可以写成形式 $Y \cdot x - X \cdot y + N = 0$，即可推知连接点 1 和点 2 的线由行列式 X，Y 和 N 确定。由此可以看出，只要比例 $X : Y : N$ 知道了，这条线也就确定了。

进一步，根据早先对线段的长度和三角形面积的考虑，X，Y 分别表示具有从点 2 指向点 1 的方向的线段在 x 轴和 y 轴上的投影，而 N 则表示三角形(0, 1, 2)取(0, 1, 2)的环行方向时的面积的两倍。显然设点 1，2 的位置改变，而保持 X，Y，N 不变，这只能是线段(1, 2)沿着所在直线移动并保持其方向和长度不变。这样就证明

了我的论断。格拉斯曼称这种有确定长度与方向,并位于一确定直线上的线段为一个定向线段,今天更常用矢量这个名称。如果一个线段在保持其长度和方向的前提下允许平行移动,甚至离开本身所在直线,则我们称之为矢量或自由矢量。由矩阵

$$\begin{pmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{pmatrix}$$

或由行列式 X, Y, N 确定的滑动矢量,是我们按格拉斯曼原理所考虑的第一个初等几何图形。

我马上要指出,量 X 和 Y 本身确定一个自由矢量,因为线段朝直线外平行移动而不改变它们。类似地,比例 $X:Y:N$ 等价于两个量,只确定无限长的直线而不确定其上的一个线段的长度。因此,这种自由矢量和无限直线是我们在这里遇到的辅助图形。辅助图形的推导原理将在后面讨论。

这些概念在初等静力学所研究的力学中起到十分重要的作用。传统上,它们在静力学中是自然地提出来的。只要是在平面内进行运算,我们就只涉及平面刚体系统静力学。对于几何处理来说,可以把定向线段当作与作用于此系统的力完全等价,由于物体的刚性,作用点可在作用力的方向上任意移动。旧力学中对力的表示法是:一条绳拴在点 2 上,一个给定的拉力大小由线段 12 来测量(图 11.2)。根据我的回忆,在旧力学书中总有一幅手拉绳子的插图,这种生动的思想方法与抽象的现代表示方法成为一个鲜明的对照①。有向线段的坐标(X, Y, N)的前两个,称为力的分量,而 N 为绕点 O 旋转的力矩。因为从直线方程可知,从点 O 到直线

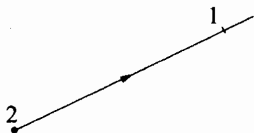


图 11.2

① 例如瓦里农(Varignon)的表:*Nouvelle Mécanique ou Statique*,巴黎,1775 年。

的垂线长为 $p=\dfrac{N}{\sqrt{x^2+y^2}}$，因而 N 实际上是距离 p 与线段的长即力的大小 $\sqrt{x^2+y^2}$ 的乘积。我们可以将这 3 个量一起作为力的坐标。解析定义在每一种情况下都可以对这些量给出非常确定的符号（这是特别重要的），就像以前几何解释一样。事实上，必须指出，为了保持公式的对称性，这里所用的方法与力学中确定力矩符号的习惯方法有所不同。习惯上是用起点 2 的坐标和自由矢量的两个坐标 (X, Y) 的行列式

$$\begin{vmatrix} x_2 & y_2 \\ X & Y \end{vmatrix} = \begin{vmatrix} x_2 & y_2 \\ x_1-x_2 & y_1-y_2 \end{vmatrix}。$$

这显然与我们的 N 数值相等而符号相反。但这个小差异，只要知道了是不会造成混乱的。

刚体力学的第一个问题是求任意力 (x_i, y_i, z_i) $(i=1, 2, \cdots, n)$ 系统的合力。合力的解析表示式为具有坐标

$$\sum_{i=1}^{n} X_i,\quad \sum_{i=1}^{n} Y_i,\quad \sum_{i=1}^{n} N_i$$

的滑动矢量。对这个问题的几何解，在图解静力学中提出了一些漂亮的方法。对两个力，干脆用著名的平行四边形法则。对 $n>2$ 的情形，则必须涉及力的多边形。一般来说，我们会找到一个唯一的滑动矢量作为任意力系的合力。然而却存在例外，例如两个大小相等、方向相反的，在不同直线上的平行力 (X, Y, N_1) 和 $(-X, -Y, N_2)$ $(N_1 \neq N_2)$ 组成的系统。合力具有分量 $(0, 0, N_1+N_2)$，显然它不能作为矢量的坐标。初等的表示法对这种现象是无能为力的，而必须把这种总是将定理的简单性和一般性打乱的不可化简的量考虑在内，称之为力偶。但如果把我们原先的公式形式地应用于矢量 $(0, 0, N_1+N_2)$，令 $\sqrt{0^2+0^2}=0$ 作为合力的大小，而

$$p=\frac{N_1+N_2}{0}=\infty$$

作为它到原点的距离，就可以很容易地把这些明显的例外纳入我们的系统。因此，如果在通常力的情况下，令其到原点的距离 p 逐渐趋向无穷，而令其大小 $\sqrt{x^2+y^2}$ 趋向于 0，使得表示转动力矩的 $p\sqrt{x^2+y^2}$ 保持不变，那么分量就正好取了这种例外的值。所以，可以把力偶的合力 $(0, 0, N_1+N_2)$ 看成是具有有限转动力矩的一个无穷小但无穷远的力。这种假想的说法是促进科学的十分方便而有用的说法，也完全符合几何中无穷远元素的传统介绍法。不管怎样，可以在力的概念引申的基础上，阐明这样一个极为一般的几何定理：在任何情况下，作用在平面上的任意一个力都具有一个合力。而按初等表示法，始终必须引出力偶的概念。

现在要研究我们的基本量在直角坐标变换下的情况，以结束我们的讨论。这将为格拉斯曼系统的具体应用提供一个有价值的分类原则。

坐标变换公式，即在直角坐标系的 4 个基本变换下，用点的原来坐标 (x, y) 表示点对新坐标系的坐标 (x', y')，其公式分别为

(1) 平移

$$\begin{cases} x' = x + a, \\ y' = y + b; \end{cases} \tag{A_1}$$

(2) 经过角 ϕ 的旋转

$$\begin{cases} x' = x\cos\phi + y\sin\phi, \\ y' = -x\sin\phi + y\cos\phi; \end{cases} \tag{A_2}$$

(3) 关于 x 轴的反射

$$\begin{cases} x' = x, \\ y' = -y; \end{cases} \tag{A_3}$$

(4) 度量单位的改变

$$\begin{cases} x' = \lambda x, \\ y' = \lambda y。 \end{cases} \tag{A_4}$$

如果对所有参数值 a, b, ϕ, λ 将这 4 种变换彼此复合起来，就得到从一个直角坐标系到另一个同时改变了单位的坐标系的最一般变换的方程组。所有可能平移与旋转的复合，对应于坐标系在平面内普通运动的总和。这些变换的总和组成一个群，即其中任何两个变换的复合，仍为该总和中的一个变换，任何变换之逆也是如此。能组成所有其他变换的特殊变换组(A)，称为群的生成元素。

在讨论这些特殊变换怎样改变我们的行列式 X, Y, N 之前，将阐述我总是予以强调，并已在这些基本几何讨论中予以突出的两个基本原理。虽然这些原理太一般了，起初有些难懂，但加以具体说明之后，会立刻变得清楚。其一是：任何图像的几何性质，在坐标系改变的情况下，必须能用不变的公式来表达，即当图形的所有点同时受我们的变换之一作用时，公式仍然保持不变。反之，任何一个公式，如果在这些坐标变换群作用下仍为不变式，则必然表示一个几何性质。现举出最简单的、大家都知道的例子，即两点之间的距离和两直线间的夹角。下面几页中，我将不得不一再涉及这些公式和诸如此类的很多公式。为了清楚起见，我提出一个非不变式的小例子：方程 $y = 0$。平面上满足此方程的点(x, y)位于 x 轴上。显然这对图形性质是完全非本质的，只对描述图形有用。类似地，每一个非不变式方程所表示的是图形与任意附加的外部东西所产生的某些关系，特别是对于坐标系所产生的某些关系，它不表示图形的任何几何性质。

第二个原理涉及由点 1, 2, …的坐标(如我们的 X, Y 和 N)组成的解析量系统。如果在坐标变换下，这个系统具有以确定方式变换成自身的性质，即如果由点 1, 2, …的新坐标形成的量的系统，完全由旧坐标以同样方法形成的这些量表达出来(旧坐标不以显式出现)，那么我们说，这个系统确定一个新的几何图形，即一种与坐标无

关的结构。事实上，我们将根据它们在坐标变换下的特性来对所有解析表达式分类，并把以同样方式变换的两个表达式系列看成是几何上等价的。

现在利用格拉斯曼基本量所提供的数据，把这些弄清楚。为此，对两点(x_1, y_1)，(x_2, y_2)进行同样的坐标变换。

(1) 我们从平移 A_1 开始

$$x_1' = x_1 + a,\quad x_2' = x_2 + a,$$
$$y_1' = y_1 + b,\quad y_2' = y_2 + b,$$

比较变换前后的矢量坐标：

$$X = x_1 - x_2,\qquad Y = y_1 - y_2,\qquad N = x_1 y_2 - x_2 y_1,$$
$$X' = x_1' - x_2',\qquad Y' = y_1' - y_2',\qquad N' = x_1' y_2' - x_2' y_1',$$

立即推知

$$\begin{cases} X' = X, \\ Y' = Y, \\ N' = N + bX - aY。\end{cases} \tag{B_1}$$

用完全同样的方法，作为变换公式，我们得到

(2) 经旋转(A_2)后

$$\begin{cases} X' = X\cos\phi + Y\sin\phi, \\ Y' = -X\sin\phi + Y\cos\phi, \\ N' = N。\end{cases} \tag{B_2}$$

(3) 作反射(A_3)后

$$\begin{cases} X' = X, \\ Y' = -Y, \\ N' = -N。\end{cases} \tag{B_3}$$

(4) 改变单位长度(A_4)后

$$\begin{cases} X' = \lambda X, \\ Y' = \lambda Y, \\ N' = \lambda^2 N, \end{cases} \tag{B_4}$$

在最后一组公式(B_4)中，由于乘数因子 λ 的次数并非永远相同而使式子有差异。我们用物理学上的因次概念来表述这个差异：X 与 Y 有线性因次 1，N 有面积因次 2。

检查这 4 组公式时，我们发现由 3 个量 X, Y, N 确定的矢量(定向线段)是真正满足我们的几何量的定义。新坐标 X', Y', N' 完全由 X, Y, N 来表达。

如果自始至终只注意 N 不出现的前两个方程，我们会看出更多的情况。新坐标系中矢量的两个坐标(X', Y')完全依赖于这些坐标的原来值；特别是，它们在平移下不改变，且在其他情况下，(X, Y)对(X', Y')的关系与(x, y)对(x', y')的关系完全一样。从上面阐明的第二条原则来看，我们可以说，两个坐标 X 与 Y 确定一个与坐标系无关的几何图形。我们已经知道，这就是自由矢量。由此可发现，前面所说的系统的原则，导出了矢量图形。

下面的讨论是在同样范围内进行。因为在所有 4 个公式组里，X', Y', N'都是作为 X, Y, N 的线性齐次函数而出现的，通过方程相除，我们看到比 $X' : Y' : N'$ 只依赖于比 $X : Y : N$。因此，这些比$X : Y : N$，确定了一个与这 3 个量的实际值无关的、独立于坐标系的几何图形。我们已经看到，这个几何图形正是无限直线。

请把公式组(B)应用于一个力偶，此时

$$X = 0,\ Y = 0。$$

于是，当然有

$$X' = 0,\ Y' = 0。$$

同时在 4 组情况下分别有

$$N' = N, \tag{C_1}$$

$$N' = N, \tag{C_2}$$

$$N' = -N, \tag{C_3}$$

$$N' = \lambda^2 N。 \tag{C_4}$$

如果对变换群作用下最多改变一个因子的量使用习惯的术语——不变量，如果根据这个因子是 1 或否而分别称之为绝对不变量或相对不变量，则可将公式组(C)称为：一个力偶的旋转矩是关于平面上所有直角坐标变换的相对不变量。

请将这个说法与起初研究过的基本几何量——三角形的面积——在坐标变换下的情况

$$\Delta = \frac{1}{2}\begin{vmatrix} X_1 & Y_1 & 1 \\ X_2 & Y_2 & 1 \\ X_3 & Y_3 & 1 \end{vmatrix}$$

相比较，平移变换(A_1)不改变这个行列式，因为它只对第一列的每个元素增加 a，对第二列元素增加 b，即分别加上第三列元素的 a 倍与 b 倍。于是，我们有

$$\Delta' = \Delta。 \tag{D_1}$$

类似地，其他 3 个变换产生

$$\Delta' = \Delta, \tag{D_2}$$

$$\Delta' = -\Delta, \tag{D_3}$$

$$\Delta' = \lambda^2 \Delta。 \tag{D_4}$$

所有这些都可以从三角形的几何意义立即推出。但这些公式和(C)组完全相同：三角形面积，从而任意面积(事实上总可以表达为三角形之和)在任意坐标变换下的性质，和力偶的旋转力矩的性质完全一

样。由此,根据第二原则,可以把这两件事看成是几何上等价的,并可以解释如下:如果在平面上有一个具有旋转力矩 N 的任意力偶,且如果用任何方式确定一个面积 $\Delta = N$ 的三角形,则在所有坐标变换下均保持此等式。也就是说,能用三角形或平行四边形或任何其他平面图形的面积来表示力偶的旋转力矩而与坐标系无关。这种几何对应是怎样导出的,以后研究类似的,虽然稍为复杂,但更具有启发性的空间关系时就会明白。

下面就结束平面几何的讨论。在平面几何中,这些抽象的道理几乎简单之极。对每一个解析公式,都可以说出一个恰当的几何意义,完全一般化的解析思想因而自动地渗入几何学中。在这方面,必须再次强调一个关键的假设:对几何图形的符号,应加上适当的规定。

第十二章
格拉斯曼空间原理

我们将用前面考虑平面时所用的方式对空间作相应的研究。因此，我们从点 1，2，3，4 的坐标形成的矩阵

$$(x_1, y_1, z_1, 1), \begin{pmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \end{pmatrix}, \begin{pmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{pmatrix}, \begin{pmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{pmatrix}$$

出发。第一个矩阵的行列式代表点自身的坐标而不需作任何考虑。

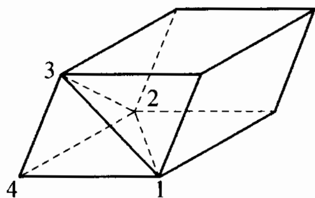


图 12.1

第四个矩阵已给出一个四阶行列式，我们已经知道，是四面体(1，2，3，4)体积的 6 倍，将其称之为“空间块”，以便与后面的术语一致。而且，可以把它干脆看做是具有棱边 41，42，43 的平行六面体(图 12.1)的体积，格拉斯曼称之为“晶体”。

第二和第三个矩阵提供了新的图形。从第二个矩阵删去两列后得到的 6 个二阶行列式为

$$\begin{cases} X = x_1 - x_2, & Y = y_1 - y_2, & Z = z_1 - z_2, \\ L = y_1 z_2 - y_2 z_1, & M = z_1 x_2 - z_2 x_1, & N = x_1 y_2 - x_2 y_1. \end{cases} \quad (1)$$

类似地，第三个矩阵给出下列 4 个三阶行列式

$$\left\{\begin{aligned} l &= \begin{vmatrix} y_1 & z_1 & 1 \\ y_2 & z_2 & 1 \\ y_3 & z_3 & 1 \end{vmatrix}, \quad m = \begin{vmatrix} z_1 & x_1 & 1 \\ z_2 & x_2 & 1 \\ z_3 & x_3 & 1 \end{vmatrix}, \\ n &= \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}, \quad B = -\begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}。\end{aligned}\right. \tag{2}$$

首先，关于 6 个行列式(1)，可以从相应的平面讨论中推知，X，Y，Z 是连接 2 到 1 的线段在各坐标轴上的投影，而 L，M，N 是三角形(0，1，2)取 0，1，2 方向在相应坐标平面上投射的面积的两倍(图 12.2)。把线段(1，2)沿它所在的直线保持长度和正向移动时，所有这些显然均保持不变。它们表示所谓空间有向线段或滑动矢量的图形。如果我们把矢量平行地移出其所在直线时，量 X，Y 和 Z 仍然保持不变，因此它们确定一个自由矢量。类似地，当有向线段在其所在直线上任意改变其长度或正向时，6 个比 $X : Y : Z : L : M : N$ 不改变，因此它们确定了无限直线。

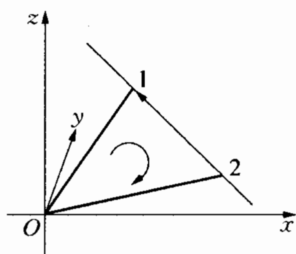


图 12.2

4 个行列式(2)首先确定了 3 个点 1，2，3 的平面。因为，我们可以写方程

$$\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} = 0$$

为如下形式

$$lx + my + nz + B = 0。$$

因此，比 $l : m : n : B$ 确定无限平面。我们进一步看到，l，m，n 是

边界方向取为 1，2，3 的三角形(1，2，3)在各坐标平面上投影面积的两倍，而 B 是四面体(0，1，2，3)体积的 6 倍，其符号对应于这些顶点的顺序。现在，当且仅当三角形(1，2，3)在它所在平面内移动和变形而其面积和边界方向不改变时，上述 4 个量显然不变。因此，它们确定了具有移动自由度的一个三角形或一个平面区域，格拉斯曼称之为“平面片”或一个“平面量值”。当我们移动平行于本身的三角形平面时，平面片的前 3 个坐标 l，m，n 也仍然不变。于是，就面积和边界方向而言，这 3 个坐标确定了在平行于本身的空间内自由移动的一个三角形，即所谓自由平面量。

如果现在回过来仔细看一下有向线段，我们首先发现，在空间中，它被 5 个参变量所决定，因为它的两个端点一共有 6 个坐标，但一个端点可沿一条直线移动。因此，上面确定的有向线段的 6 个坐标 X，Y，Z，L，M 和 N 不可能彼此独立，而必须满足一个条件。我们可以从行列式定理(这总是我们理论的关键)来推出这个条件。请考虑行列式：

$$\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \end{vmatrix} = 0。$$

由于两行元素完全相同，因而恒等于 0。将行列式展开成上两行和下两行各对应的代数余子式乘积之和。包含两个被虚线围住的代数余子式的第一个被加项正是 $N \cdot Z$，而整个行列式为

$$2(N \cdot Z + M \cdot Y + L \cdot X)。 \tag{3}$$

因而得到任意有向线段的 6 个坐标所必须满足的等式 $X \cdot L + Y \cdot M + Z \cdot N = 0$。借助公式(1)，不难证明，6 个量之间的方程(3)足以使 6 个量代表一个有向线段的坐标。我不需要对这个十分基本的问题作任何讨论。

现在再谈一谈这些概念在力学上的应用。和平面上的情形一样,现在的有向线段代表一个作用于空间刚体上的力,包括作用点、大小和方向。在有向线段的 6 个坐标中,称 X, Y, Z 为平行于坐标轴的分力,而 L, M, N 是围绕这些轴的转矩,其所选方向仍与力学中常取的方向相反。3 个分量 X, Y, Z 决定了力的大小和方向,其方向余弦是比 $X:Y:Z$。我们得到的力是平行六面体的对角线,其分力的大小是坐标轴上的线段 X, Y, Z。利用 L, M, N 作出同样的对角线,我们得到一个确定的方向,称为合力矩的轴的方向。按著名的空间几何的公式,方程(3)表示力的方向和合力矩的轴的方向互相垂直。和平面上的情况一样,我们将力偶作为 $X=Y=Z=0$,由于L, M, N不全为零的极限情形而把它包括在有向线段的概念之内。简单地过渡到极限,即表明:这里是指一个旋转力矩为有限的,作用于无限远的无穷小力。初等理论里避免这种表达方式,而认为一个力偶只不过是两个作用在不同平行直线上的,大小相等方向相反的力:(X, Y, Z, L_1, M_1, N_1)和$(-X, -Y, -Z, L_2, M_2, N_2)$。它们的和,事实上就给出刚才假设过的坐标$(0, 0, 0, L_1+L_2, M_1+M_2, N_1+N_2)$。

现在考虑作用在一个刚体上的任意力系 X_i, Y_i, Z_i, L_i, M_i, $N_i(i=1, 2, \cdots, n)$的合成。在初等教材和讲义里,花了很多时间讲这个问题,而这里却能迅速加以处理;因为初等讨论中不考虑符号规则,只有不厌其烦地考虑各个具体情况,而我们有了解析公式,就觉得多此一举了。合成的基本原则是,我们求出和

$$\Xi=\sum_{i=1}^{n}X_i,\quad H=\sum_{i=1}^{n}Y_i,\quad Z=\sum_{i=1}^{n}Z_i,$$

$$\Lambda=\sum_{i=1}^{n}L_i,\quad M=\sum_{i=1}^{n}M_i,\quad N=\sum_{i=1}^{n}N_i,$$

并把它们作为力系的坐标,或用普吕克的话,作为动力坐标。这里,再一次区分出沿坐标轴的 3 个分量和围绕坐标轴的 3 个旋转力矩。

现在，这个力系一般将不是一个单力，因为 6 个和式不一定满足单个有向线段的坐标条件

$$\Xi \cdot \Lambda + H \cdot M + Z \cdot N = 0。$$

这是在空间中所出现的、与平面上的情形相反的新概念，即一个作用于刚体的力系不一定能简化成一个简单的力。

为了获得一个力系的具体图像，将以尽可能简单的，作为尽可能少的几个力的合力的方式加以表达。我们将证明，能将每个力系考虑为一个单个力和一个轴线与该力的作用线（所谓力系的中心轴）平行的力偶的合力，且此分解是唯一的。这种作用于刚体的力的合成理论，在 1804 年出版并已一版再版的普昂索（Poinsot）的 *Eléments de statique*① 中有经典的表述。我们这里所谈的，实际是普昂索的中心轴。普昂索的处理方法是初等几何方法，是非常复杂的，现在初等数学教学中仍然如此。

为了证明上面的定理，我们要指出，从力系中抽去一个力偶可能产生的任何单力，必须具有以 Ξ, H 和 Z 作为平行于坐标轴的分量。因此，如果此力偶的转矩轴平行于中心轴，则其分量必然与 Ξ, H 和 Z 成比例。设其 6 个坐标为 0, 0, 0, $k\Xi$, kH, kZ，其中 k 为待定参量。为了从这个力偶求得我们的系统(Ξ, H, Z, Λ, M, N)，必须把系统

$$\Xi,\ H,\ Z,\ \Lambda - k\Xi,\ M - kH,\ N - kZ$$

加到力偶上。如果能确定 k 使该系统成为一个单力，此定理即得到证明。对此的必要充分条件是坐标满足方程(3)，即

$$\Xi(\Lambda - k\Xi) + H(M - kH) + Z(N - kZ) = 0。$$

由此可得唯一解

$$k = \frac{\Xi\Lambda + HM + ZN}{\Xi^2 + H^2 + Z^2}。$$

① 第 12 版，J·贝特兰(J. Bertrand)出版，巴黎，1877 年。

因为我们假设分母不为零，否则我们处理的是一个力偶而不是一个真正的力系。如果对 k 指定被普吕克称为动力参数的这个值，那么实际上就是把这个力系分解成了一个力偶和一个单力，而证明的过程已表明这个合成是唯一的。

现在的问题是：这个解可以和什么几何表示联系起来？这些研究又得回溯到莫比乌斯 1837 年发表的《静力学》一书①。在这本书里，他问到所谓"零轴"，即力系绕它的转矩为零的问题。他称所有这些零轴的系统为一个"零系统"。这个术语无疑是大家所熟悉的，其起源即在于此。

现在必须对这里应用的旋转力矩或力矩的一般概念下定义。设在空间给出两个有向线段(1, 2)和(1′, 2′)(图 12.3)，由它们构成四面体(1, 2, 1′, 2′)，其体积为

$$\frac{1}{6}\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_1' & y_1' & z_1' & 1 \\ x_2' & y_2' & z_2' & 1 \end{vmatrix}。$$

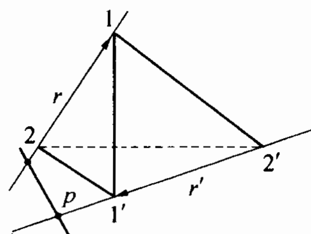


图 12.3

将此行列式展开成上两行和下两行代数余子式乘积之和，得到 $\frac{1}{6}(XL'+YM'+ZN'+LX'+MY'+NZ')$，其中 X', …, N'是有向线段(1′, 2′)的坐标。所出现的两个有向线段坐标的双线性组合

$$XL'+YM'+ZN'+LX'+MY'+NZ'$$

称为一有向线段相对于另一有向线段之矩，它等于以各有向线段的端点为顶点的四面体体积的 6 倍，因而是一个独立的几何量。如果 r 和 r'是各有向线段的长度，ϕ 是它们之间的夹角，p 是它们的公垂

① 莱比锡，1837 年，或见其《著作集》第 3 卷，莱比锡，1896 年。

线长，则由初等几何可知，在适当选择 ϕ 的符号情况下，矩等于 $r \cdot r' \cdot p \cdot \sin\phi$。

如果我们用无穷直线代替有向线段(1, 2)，则有向线段(1′, 2′)相对于直线的矩，将定义为相对于该直线上长 $r=1$ 的有向线段的矩(按前面意义)，即 $r'p\sin\phi$。这是将前面的表达式除以 $r(=\sqrt{X^2+Y^2+Z^2})$ 而得到的结果。因而，有向线段 (X', Y', Z', L', M', N') 相对于包含有向段线 (X, Y, Z, L, M, N) 的无限直线的矩为

$$\frac{XL'+YM'+ZN'+LX'+MY'+NZ'}{\sqrt{X^2+Y^2+Z^2}}。$$

事实上，这个值只依赖于 6 个量 $X, \cdots, N$ 之比加上一个为它们所共有的符号。因而，当无限直线及其上的一个正向已知时，它就完全被确定了。这个矩正是静力学中所谓由一个有向线段代表的力绕作为轴的直线的旋转矩，虽然所选的符号有所不同。

现在考虑一个动力系统的力系

$$\Xi=\sum_{i=1}^{n} X_i', \cdots, N=\sum_{i=1}^{n} N_i'$$

的矩或旋转矩。这自然指的是若干力的矩之和，即表达式

$$\sum_{i=1}^{n} \frac{XL_i'+YM_i'+ZN_i'+LX_i'+MY_i'+NZ_i'}{\sqrt{X^2+Y^2+Z^2}}$$
$$=\frac{X\Lambda+YM+ZN+L\Xi+MH+NZ}{\sqrt{X^2+Y^2+Z^2}}。$$

如果在这个表达式中，我们依次令无限直线 $X, \cdots, N$ 与 3 个正坐标轴相等，由此表达式依次取值 Λ, M, N，这也表明了前面对这些量指定的记号是合理的。

现在可以着手处理莫比乌斯提出的问题：如果 $\Lambda X+MY+NZ+\Xi L+HM+ZN=0$，则给出的力系 $\Xi, H, \cdots, N$ 相对直线 $(X:Y:\cdots:N)$(零轴)的矩为 0。因此，动力系统的零系统是由这

个方程给出的全部直线($X:Y:\cdots:N$)。因为系统 Λ, …, Z 作为一个动力系统的坐标可以是 6 个自由的量,所以这是 6 个量 X, …, N 的最一般的线性齐次方程。普吕克和莫比乌斯一起,作为 19 世纪解析几何领域里的先锋,正是研究了这些由任意线性齐次方程确定的全部直线,并称之为线丛(此后将详细讨论)。因此,莫比乌斯的零系统,就正是普吕克的线丛。

现在力图对这个零系统给出一个尽可能清晰的图像。当然,我们不能说出这个术语在真正意义上的一个几何图像,因为空间中充斥着无限繁多的零系统。然而,它的分类是不难理解的。为此,按照这些讲稿中所一贯遵循的计划,将尽可能方便地选择坐标轴,这里就选动力系统的中心轴为 Z 轴。因为,正如我们所知道的,动力系统是沿中心轴作用的一个力和转轴平行于中心轴的一个力偶的合力,按我们对 Z 轴的选择,4 个坐标 Ξ, H, Λ, M 必然为 0,故 Z 代表了该力的度量,N 则表示该力偶绕其轴的转矩。因此,此动力系统的参数为

$$k=\frac{\Xi\Lambda+HM+ZN}{\Xi^2+H^2+Z^2}=\frac{N}{Z}。$$

于是,新坐标系中的线丛方程具有简化形式

$$NZ+ZN=0,$$

或在除以 Z 后为

$$k\cdot Z+N=0。\tag{4}$$

这个形式就当作以后讨论的基础。如果 $P_1(x_1, y_1, z_1)$ 和 $P_2(x_2, y_2, z_2)$是零系统的一条直线($X:Y:Z:L:M:N$)上的两个点,那么因 $Z=z_1-z_2$, $N=x_1y_2-x_2y_1$,对零线上任意两点的坐标,方程(4)给出条件

$$k(z_1-z_2)+(x_1y_2-x_2y_1)=0。\tag{5}$$

现在如果使 P_2 固定，则方程(5)是零系统中一条线上所有点 P_1 的坐标(x_1, y_1, z_1)的方程，而且 P_2 也在这条直线上。为了清楚起见，如果用流动坐标(x, y, z)代替(x_1, y_1, z_1)，我们看到，所有点 P_1 填满一个平面，其方程为

$$y_2x - x_2y + k \cdot z = kz_2。\tag{5'}$$

本平面包含点 P_2，因为 $x = x_2$，$y = y_2$，$z = z_2$ 满足此方程。由此证明，通过空间任意点 P_2 有无穷多个零线，它们形成一个填满平面(5′)的平面线束。如果能得到对应每点 P_2 的这个平面(零平面)位置的清楚图像，我们的问题就解决了。

在空间中作平行于 Z 轴平移和绕 Z 轴旋转的变换情况下，在方程(5)内出现的两个表达式 $N = x_1y_2 - x_2y_1$，$Z = z_1 - z_2$ 具有不变的性质，因为平移使 X 和 Y，从而使 N 以及差 $z_1 - z_2$ 都保持不变，而旋转则不影响 z 坐标(即 Z)并使 xy 平面内的面积 N 保持不变。于是，在空间绕中心轴——在此为 z 轴——按顺时针方向旋转和沿该轴平移时，方程(5)及其确定的零系统变为自身。

这个定理使我们的问题变得相当容易了。只要知道零系统中哪个平面属于 X 轴的正半轴上的任意一点，自然也就知道了属于空间每个点的零平面。因为通过沿 Z 轴的平移和绕 Z 轴的旋转，可以使 X 轴上一点与空间的任一定点重合，而据我们的定理，对应的零平面变为自身。换句话说，垂直于中心轴的一条半轴，其各点的零平面，有一个相对于半轴与中心轴的位置，这个位置与半轴的选择无关。

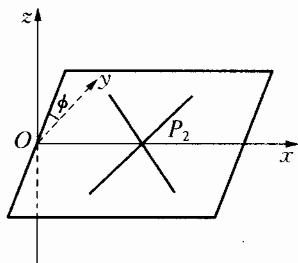


图 12.4

如以 X 轴为限，令 $y_2 = z_2 = 0$，由方程(5)得到 $kz - x_2y = 0$。因为 $y = z = 0$ 恒满足此方程，所以它通过 X 轴(图 12.4)。如改写方程为形式 $\frac{z}{y} =$

$\frac{x_2}{k}$，推得平面对于水平（xy 平面）面的倾角的正切为

$$\tan\phi = \frac{x_2}{k}。$$

于是，平面的位置就完全被确定。在图 12.5 中，画出了它在与其垂直的 yz 平面上的截痕。

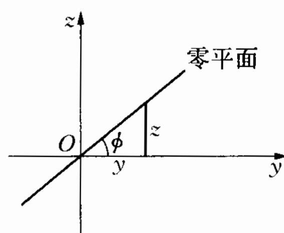


图 12.5

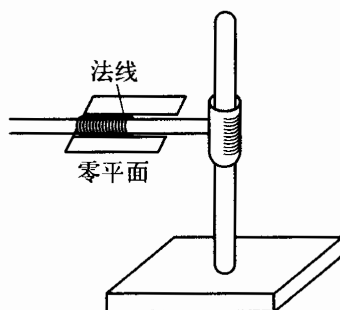


图 12.6

综上所述，我们可以说，这个结论与空间坐标系的选择无关。对于看做垂直的离中心轴有 r 那么远的每一个点，有一个包含从这点到轴的垂线的、属于零系统的平面，它对水平面的倾角的正切为 $\frac{r}{k}$。如果沿垂直于此轴的半射线移动此点，则相应的零系统的平面在 $r=0$ 时为水平面，随 r 增加将按 $k>0$ 或 $k<0$ 而上倾或下倾，当 r 变成无穷时渐近趋向于垂直。我们可以用席林的模型（图 12.6）把这些关系讲清楚。在这个模型中，有一个可沿中心轴滑动与绕中心轴旋转的滑臂，臂上附有一个平面板，当板从轴上离远时，以适当的方式升高。

现在来考虑点 P_2 所经过平面的法线方向。它的方向余弦比与平面方程（5′）的系数之比，即 $y_2:(-x_2):k$ 相同。

可以把这个方向看做是在空间作无穷小螺旋运动时点 P_2 的移动方向。事实上，如果使空间像一个刚体那样绕 Z 轴转过有限角 ω

并同时平行于 Z 轴移动量 c，每个点(x, y, z)将变到由方程

$$\begin{cases} x' = x\cos\omega - y\sin\omega, \\ y' = x\sin\omega + y\cos\omega, \\ z' = z + c \end{cases}$$

确定的新位置。我们用$-\mathrm{d}\omega$代替 ω，并令 $c = k\mathrm{d}\omega$，把有限的螺旋运动过渡到无穷小的螺旋运动。负号表示如果沿 Z 的正向移动，则 $k > 0$ 时 xy 平面的旋转是负的，即螺旋运动是负的（左手螺旋）。忽略 $\mathrm{d}\omega$的二阶和高阶量，即令 $\cos\mathrm{d}\omega = 1$，$\sin\mathrm{d}\omega = \mathrm{d}\omega$，我们得

$$x' = x + y\mathrm{d}\omega,\ y' = -x\mathrm{d}\omega + y,\ z' = z + k\mathrm{d}\omega。$$

在此无穷小的螺旋运动下，给定点 P_2 的坐标的增量为 $\mathrm{d}x_2 = y_2\mathrm{d}\omega$，$\mathrm{d}y_2 = -x_2\mathrm{d}\omega$，$\mathrm{d}z_2 = k\mathrm{d}\omega$，即点 P_2 将沿方向

$$\mathrm{d}x_2 : \mathrm{d}y_2 : \mathrm{d}z_2 = y_2 : (-x_2) : k \qquad (6)^{①}$$

移动。事实上，这正是沿法线(6)的方向。因此，如果我们使空间绕中心轴作一个无穷小螺旋运动，使得沿轴的移动等于转角（取负号）的 k 倍，则属于空间任意点的、参数为 k 的零系统中的平面，将垂直于该点移动的弧线。

因为螺旋运动的表示是十分容易的，所以对各平面在零系统中的分布，可以用这个方法得到一个生动的图像。例如，增加点到中心轴的距离 r，在螺旋运动中所经路径的水平投影 $r\mathrm{d}\omega$ 增长，路径本身就越平，因为增加高度 $k\mathrm{d}\omega$ 是常量，又因增高方向与路径相垂直，所以零系统中的平面就越陡。如果将无限次的这种无穷小螺旋运动复合成一个空间内的连续螺旋进动，则距中心轴为 r 的每个点将画出一条对水平的倾角斜率为 $-\frac{k}{r}$、螺距为与 r 无关的

① 式码为编辑所加。

$2\pi k$ 的螺旋线。与此螺旋线相垂直的诸平面都是零系统中的平面。

讨论了零系统中的平面之后，我现在尽量为零轴描出一个清楚的图像。任取一零轴 g（图 12.7），并作 g 与中心轴的一条公垂线，分别与 g 及中心轴相交于点 P 和点 O。这样，作为从点 P 到中心轴的垂线的 PO 属于零系统，OPg 必然是属于 P 的零系统的平面。因 g 垂直于 OP，它与水平面的交角与零平面和水平面的交角一样为 ϕ，即 $\tan\phi = \frac{r}{k}$，$r = OP$。因此，如果通过垂直于中心轴的每条半射线上的每个点 P，都作一条直线与此射线相垂直，并使其与水平面夹角的正切 $\tan\phi = \frac{r}{k}$（其中 r 为点 P 到中心轴的距离），就得到了所有的零轴。

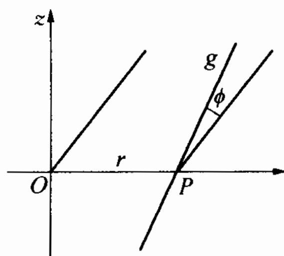


图 12.7

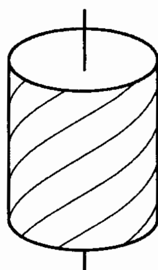


图 12.8

可以把这个结构说得更清楚一些。取一个半径为 r、轴为中心轴的圆柱面（图 12.8），在上面画出与水平面的倾角 ϕ 的正切为 $\tan\phi = \frac{r}{k}$ 的所有螺旋线。这些螺旋线的切线的总体，显然与距离中心轴为 r 的所有零轴相同。通过改变 r，就得到所有零轴。向外移动时，这些螺旋线变陡。它们在每一点上以对应的零平面为其密切平面。因此，它们与刚才说的螺旋线成直角，而上述螺旋线在每一点上都是与零平面垂直的。

讨论了螺旋线和零系统之间呈现的双重联系之后，就可以理解为什么全部理论与螺旋线联系在一起。罗伯特·波尔(Robert Ball)爵士在他著的《螺旋理论》①中就曾利用这个指定，在书中讨论了与作用于刚体的力系相联系的一切几何关系。

现在把我们所讲的系统总结一下。借助格拉斯曼原则，获得了4个基本几何图形：点、线段、平面片和空间块。像在平面上的情形一样，我们将讨论这些图形在直角坐标变换下的性质，并按一般的原则加以分类。

① 罗伯特·波尔，《螺旋理论》，都柏林，1876年。

第十三章

直角坐标变换下空间基本图形的分类

首先，当然应对空间中所有直角坐标系变换有一个概念。这些变换是所有空间几何的基础，所以在本讲义中无论如何不能对其忽略。和平面上的情况一样，要考虑的最一般的坐标系的变换，由下列各部分组成：

(1) 平移；

(2) 绕原点旋转；

(3) 反射；

(4) 长度单位的改变。

平移的方程当然就是

$$\begin{cases} x' = x + a, \\ y' = y + b, \\ z' = z + c。\end{cases} \tag{A_1}$$

在任意情况下的旋转方程具有形式：

$$\begin{cases} x' = a_1 x + b_1 y + c_1 z, \\ y' = a_2 x + b_2 y + c_2 z, \\ z' = a_3 x + b_3 y + c_3 z。\end{cases} \tag{A_2}$$

我们将考虑系数行列式，这里比平面上的情况要复杂一些。这两类变换的所有可能的复合，产生构成空间坐标系的一切适当运动。

正如平面上对一个轴反射一样，这里可以考虑对一个坐标平面，

例如 xy 平面的反射，并得

$$x' = x,\ y' = y,\ z' = -z,$$

但能借助 3 个负号把这些公式写成更对称的形式

$$x' = -x,\ y' = -y,\ z' = -z。\tag{A_3}$$

这是一个对原点的反射，有时称为反演①。在平面上

$$x' = -x,\ y' = -y$$

就不是一个反射，而是扫过 180°的旋转。一般来说，只在奇数维空间，对原点的反演才是一个反射，对偶维空间则是旋转。

最后，单位长度的改变由方程

$$x' = \lambda x,\ y' = \lambda y,\ z' = \lambda z \quad (\lambda > 0)\tag{A_4}$$

给出。如 $\lambda < 0$，则此变换为一个反射加单位长度的改变。

剩下需要较仔细地考虑旋转公式。正如你们所知道的，绕原点的最一般的旋转依赖于 3 个参数，因为轴旋转的方向余弦代表两个独立量；此外，旋转角是任意的。四元数理论对依赖 3 个独立参数的所有旋转提供了对称的处理。不过，欧拉在四元数发现之前已为所述问题建立了公式。这里将介绍力学教科书里通常的处理方法，这个方法利用了新轴相对于旧轴的 9 个方向余弦。我们从前面给出的公式

$$\begin{cases} x' = a_1x + b_1y + c_1z, \\ y' = a_2x + b_2y + c_2z, \\ z' = a_3x + b_3y + c_3z \end{cases}\tag{1}$$

出发。请考虑旧 x 轴上的一点 x，$y = 0$，$z = 0$。它的新坐标为 $x' = a_1x$，$y' = a_2x$，$z' = a_3x$，即 a_1，a_2，a_3 是旧 x 轴与各新轴交角的余

① 有时“反演”用于以半径倒数进行的完全不同的变换。

弦。类似地，b_1，b_2，b_3 和 c_1，c_2，c_3 分别是旧 y 轴、旧 z 轴和各新轴交角的余弦。

变换方程的这 9 个系数并不是相互独立的。我们可以从刚才的说明中推导出它们之间的关系，或利用在每个正交代换中已知的关系，即原点固定时在每一个旋转和反射中有

$$x'^2+y'^2+z'^2=x^2+y^2+z^2。\tag{2}$$

它表示到点 O 的距离是不变量。我们选择第二个方法。

(a) 将(1)式代入(2)式，比较相应项系数，可得 9 个量 a_1，…，c_3 间的下列 6 个关系式

$$\begin{cases}a_1^2+a_2^2+a_3^2=1, & b_1c_1+b_2c_2+b_3c_3=0,\\ b_1^2+b_2^2+b_3^2=1, & c_1a_1+c_2a_2+c_3a_3=0,\\ c_1^2+c_2^2+c_3^2=1, & a_1b_1+a_2b_2+a_3b_3=0。\end{cases}\tag{3}$$

(b) 用 3 个量 a，b，c 分别乘以方程组(1)并利用(3)式求解，可得

$$\begin{cases}x=a_1x'+a_2y'+a_3z',\\ y=b_1x'+b_2y'+b_3z',\\ z=c_1x'+c_2y'+c_3z'。\end{cases}\tag{4}$$

显然，这是从方程组(1)中将其系数矩阵的行列互换而得到的所谓转置线性变换。

(c) 另一方面，用行列式法则解方程(1)，可得

$$x=\frac{1}{\Delta}\begin{vmatrix}x' & b_1 & c_1\\ y' & b_2 & c_2\\ z' & b_3 & c_3\end{vmatrix},\cdots,\text{其中 }\Delta=\begin{vmatrix}a_1 & b_1 & c_1\\ a_2 & b_2 & c_2\\ a_3 & b_3 & c_3\end{vmatrix}。$$

此处 x' 的系数必然与方程组(4)的第一个方程一样，即

$$\frac{1}{\Delta}\begin{vmatrix}b_2 & c_2\\ b_3 & c_3\end{vmatrix}=a_1。\tag{5}$$

对其他系数也有类似的结果，即正交变换的每一个系数必然等于系数矩阵的对应代数余子式除以行列式 Δ。

(d) 现在计算行列式 Δ。为此，利用行列式的乘法规律求它的平方：

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \cdot \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

$$= \begin{vmatrix} a_1^2+a_2^2+a_3^2 & b_1a_1+b_2a_2+b_3a_3 & c_1a_1+c_2a_2+c_3a_3 \\ a_1b_1+a_2b_2+a_3b_3 & b_1^2+b_2^2+b_3^2 & c_1b_1+c_2b_2+c_3b_3 \\ a_1c_1+a_2c_2+a_3c_3 & b_1c_1+b_2c_2+b_3c_3 & c_1^2+c_2^2+c_3^2 \end{vmatrix}。$$

根据(3)式，此积为

$$\Delta^2 = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1,$$

所以 $\Delta = \pm 1$。为了决定选择什么符号，我们要指出，至今为止，我们只用到关系(2)，它对于旋转和反射是同样成立的，在所有正交变换中，旋转具有从恒等变换 $x' = x,\ y' = y,\ z' = z$ 通过连续改变系数而产生的性质，它相应于一个坐标系的连续运动，从原位置变到新位置。另一方面，一般称为反射的变换，是从反演 $x' = -x$，$y' = -y,\ z' = -z$ 的连续变形而产生的，而反演本身不能从恒等变换连续地产生。但变换的行列式是系数的连续函数，当我们连续地将恒等变换改变成任意旋转时，必然是连续地变化。其开始值为

$$\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = +1。$$

因为它的值只能是 $+1$ 或 -1，但对于旋转，它必须保持为 $+1$，突然改变到 -1 意味着不连续性。因此，对每一个旋转，行列式的值为

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = +1。\tag{6}$$

对于反射，则必然有 $\Delta = -1$。

现在，(5)式具有简单的形式

$$a_1 = \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix},\tag{7}$$

因此，直角坐标系的旋转变换矩阵的每一个系数都等于对应的代数余子式。

现在转入实际问题，在直角坐标系的 4 个类型的变换下，如何求出直线段 X, Y, Z, L, M, N，平面片 l, m, n, D 及空间块 T 这些空间图形的坐标变化。

把所有这些公式都写出来要占用许多篇幅，而且不厌其烦。因此，我只说明值得特别注意的几点。首先，我指出几个你们很容易验明的性质，在所有直线段坐标的变换公式中，新系的前 3 个坐标 X', Y', Z' 只通过 X, Y, Z 表示，事实上是它们的线性齐次函数。量 L, M, N 不在表示式内。因此，根据前面说的一般原则，3 个量 X, Y, Z 一起必然确定一个与坐标系无关的几何图形，这就是我们说过的自由矢量。同样，平面片的 3 个坐标 l, m, n 的变换与第四个坐标无关，因而它们也具有独立于坐标系的几何意义。它们代表前面说过的自由平面度量。

现在通过专门的计算，找出在坐标变换(A_1), …, (A_4)下自由矢量坐标 X, Y, Z 的性质。为此，由公式(A_2)，只需用 x, y, z 代替 $X' = x_1' - x_2'$, …中的 x_1'…，立即得到下列公式：

(1) 平移

$$X' = X,\ Y' = Y,\ Z' = Z。\tag{B_1}$$

（2）旋转

$$\begin{cases} X' = a_1 X + b_1 Y + c_1 Z, \\ Y' = a_2 X + b_2 Y + c_2 Z, \\ Z' = a_3 X + b_3 Y + c_3 Z。\end{cases}\tag{B_2}$$

（3）反演

$$X' = -X,\ Y' = -Y,\ Z' = -Z。\tag{B_3}$$

（4）长度单位的改变

$$X' = \lambda X,\ Y' = \lambda Y,\ Z' = \lambda Z。\tag{B_4}$$

因此，在坐标系平移时，自由矢量的坐标不变；但对其他变换，它们的变换公式和点坐标的一样。

请与力偶的变换公式作比较，在线段的坐标变换公式中令 $X = Y = Z = 0$，即得力偶变换公式。于是，当然有

$$X' = Y' = Z' = 0。$$

而对关于新轴的旋转力矩，我们有下列公式：

（1）平移

$$L' = L,\ M' = M,\ N' = N。\tag{C_1}$$

（2）旋转

$$\begin{cases} L' = a_1 L + b_1 M + c_1 N, \\ M' = a_2 L + b_2 M + c_2 N, \\ N' = a_3 L + b_3 M + c_3 N。\end{cases}\tag{C_2}$$

（3）反演

$$L' = L,\ M' = M,\ N' = N。\tag{C_3}$$

(4) 长度单位的变化

$$L' = \lambda^2 L, M' = \lambda^2 M, N' = \lambda^2 N。\qquad (C_4)$$

力偶的坐标在坐标系的平移和反演下不改变，在旋转时与点坐标性质相同，当单位长度改变时它们乘以因子 λ^2，即它们有因次 2(平面的因次)，而自由矢量和点坐标一样具有因次 1。

导出公式(C_1)，(C_3)，(C_4)毫无困难；或许对(C_2)需作某些解释。事实上，借助于公式(A_2)，有

$$L' = \begin{vmatrix} Y_1' & Z_1' \\ Y_2' & Z_2' \end{vmatrix} = \begin{vmatrix} a_2x_1 + b_2y_1 + c_2z_1 & a_3x_1 + b_3y_1 + c_3z_1 \\ a_2x_2 + b_2y_2 + c_2z_2 & a_3x_2 + b_3y_2 + c_3z_2 \end{vmatrix}。$$

如果将后一个行列式乘开，我们得到 $3 \times 3 + 3 \times 3 = 18$ 项，其中有 3 组，每组两项相抵消(例如 $a_2x_1 \cdot a_3x_2 - a_3x_1 \cdot a_2x_2$，…)，剩下 12 项可组合成下列行列式乘积之和：

$$L' = \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} \cdot \begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix} + \begin{vmatrix} c_2 & a_2 \\ c_3 & a_3 \end{vmatrix} \cdot \begin{vmatrix} z_1 & x_1 \\ z_2 & x_2 \end{vmatrix} + \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \cdot \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}。$$

根据(7)式，前一个因式分别等于 a_1，b_1，c_1，而后一个因式是 L，M，N。因而求出上面 L' 的公式。类似地，可得 M'，N' 的另两个公式。

作为第三个图形，请考虑自由平面度量。与上面的计算相类似，这个计算十分简单，留给你们去完成。计算将证明，自由平面度量的分量 l，m，n 在所有情况下和力偶的坐标 L，M，N 的变化情况一样。

为了清楚起见，把这些结果列入下页一个小表，它给出第一个坐标的变换，其他坐标可通过循环交替求得。

	平 移	旋 转	反 演	单位长度改变
自由矢量	X	$a_1X+b_1Y+c_1Z$	$-X$	λX
力 偶	L	$a_1L+b_1M+c_1N$	L	λ^2L
自由平面度量	l	$a_1l+b_1m+c_1n$	l	λ^2l

现在我们得到一系列几何叙述的精确基础，这些在教科书中往往不出现，或只是偶尔出现，而在其表达形式中，几何意义是不明显的。这里考虑的几何图形往往并不像我们所认为必须的那样，分析得一清二楚，结果有趣的关系完全模糊不清了。例如，甚至对普昂索，力偶和自由平面度量的概念从一开始就总是混在一起的。显然，这就难以理解。对我们而言，根据前述一般原则，把上面表中最后两行作比较表明，一个力偶和一个自由平面度量要看做是同类型的基本几何图形，因为在所有直角坐标系的变换下它们的性质是相同的。请把这段说明的含意弄清楚。如给定一个力偶 L, M, N 及一个平面度量 l, m, n，用方程 $l=L, m=M, n=N$（或者反过来从 l, m, n 出发）建立起它们之间的关系，则在任何坐标变换下上列等式不受影响。因此，可不利用坐标系，作纯几何描述。为此，从平面度量 l, m, n 出发，选取方便的坐标系，使得 $l=m=0$。这样，自由平面度量代表在 xy 平面上或在与 xy 平面平行的平面上的一个三角形(1, 2, 3)，且 n 是它的面积的两倍，即等于平行四边形(1, 1′, 2, 3)的面积，其符号则由环行方向 11′2(图 13.1)决定。现在我断言，力矩为 $L=M=0$, $N=n$ 的相应的力偶，可以由平行四边形的对边(1, 1′)和(2, 3)，以及 1 和 2 之间箭头所指的力形成。为证明起见，我仍在 xy 平面上选择一个方便的坐标系，即以直线 11′为 y 轴而 x 轴则通过

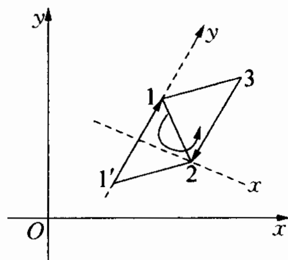


图 13.1

点 2(图 13.1 中的虚线)。于是,两个有向线段$(1, 1')$和$(2, 3)$及其所形成的力偶,有旋转力矩 $L = M = 0$。而且,线段$(1, 1')$的第三个旋转力矩也是零,因而 N 等于$(2, 3)$的旋转矩

$$N = \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} = x_2 \cdot y_3$$

(因按我们的选择,$x_2 = x_3$,$y_2 = 0$)。另一方面,对坐标系的这种位置,平面的第三个坐标是

$$n = \begin{vmatrix} 0 & y_1 & 1 \\ x_2 & 0 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = x_2 \cdot y_3,$$

即平行四边形的底 y_3 与高 x_2 的乘积。因此,不论从符号还是大小方面来说,都有 $N = n$。这证明了我的论断。

我们可以把这个当作与坐标系无关的一般的结果:一个有确定周界方向的平行四边形所代表的自由平面及由此平行四边形两对边按原给方向相反的箭头指向给出的力偶是几何等价图形,即对每一个坐标系,它们的分量相同。因此,由此定理,任何时候都可用平行四边形代替力偶,或相反。

对前述表中的第二行,不必作进一步的讨论,我们将比较第一、第三行,即自由矢量与自由平面度量。首先指出,两者在平移与旋转下有同样的性质,但加上反射与长度单位的改变,差异就出现了。为了详细讨论起见,我们设想在熟悉的(右手)坐标系内给出一个平面度量 l, m, n,并用方程 $X = l, Y = m, Z = n$ 使它同一个自由矢量联系起来。如果以坐标系的运动为限,则这些方程将保持不变,但由于反射和长度单位的改变,它们将发生变化。如果希望对它们给出几何表示,不能不考虑坐标系的方向和长度单位。事实上,如果还是选择前面的坐标系,使得 $l = m = 0$,而 n 等于在 xy 平面上的平行四边形$(1, 1', 2, 3)$的面积,则图形(图 13.2)表明 $n > 0$,且矢量 X

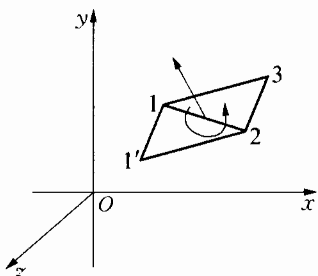


图 13.2

$=Y=0$，$Z=n$ 具有 z 轴的正方向。显然，可以把这个与坐标系的位置无关的事实叙述如下：为了在一个右手坐标系内获得一个与给定平面度量的坐标相同的自由矢量，我们对平面作一条法线，从正向看下去，表示平面度量的平行四边形的周界沿逆时针方向旋转，并在平面上画出一个等于平行四边形面积的线段。这样，不管坐标系发生平移或旋转，这个矢量与平面度量之间的等价性保持不变。但如果使坐标系反演或改变单位长度，则不再相等。例如，用分米代表厘米测量时，面积度量要除以 100，而矢量段只除以 10。在反演情况下，矢量改变符号，而平面度量则不变。

只有在选定了坐标系的方向和长度单位而不再改变的前提下，才能使一个自由平面度量与一个自由矢量完全等同。当然，每一个人可随意加上这种限制，但如果要使别人了解的话，他必须认识到他的选择的任意性。所有这些事情都是非常清楚简单的，但必须记住，因为由于历史演变的原因，在今天的物理学上留下了某种含混的地方。因此，有必要谈谈这一段历史。

我前面强调过，1844 年发表的格拉斯曼的延拓理论，由于非常难读，对物理学界与力学界没有留下什么印象。大约在同一时期，W·R·哈密顿在都柏林的研究，却在英格兰产生了较大的影响。哈密顿是四元数的发明者[①]，也是我们称为自由矢量的“矢量”这个词的引入者，不过他没有使用“滑动矢量”这个概念，而且没有区分自由平面度量和自由矢量，因为他一开始就假设坐标系方向和长度单位是确定的。这种习惯说法渗入到物理学后，使得长时间内对真实

① 见第一卷第一部分第四章第一节，我曾在那一节进行了详细的讨论。

矢量和平面度量不加区分。但在深入地研究中，已逐渐产生了按在反演变换下的性质把统称为矢量的两种含义加以区分的要求。为此引入"极性"、"轴性"等词。一个极性矢量在反演时改变其符号，因此与我们的自由矢量相同；一个轴性矢量在反演时不改变其符号，因此与我们的自由平面度量相同（这里不考虑因次）。最终，物理学不得不承认这里存在着差别，它在某种程度上使人惊奇，但在一般叙述中还是出现了区分。不过在我们的一般处理中，从一开始就很自然地加以区分了。

我们举一个例子，以澄清这个模糊的概念。电激发是一个极性矢量的说法，表示它能用 3 个量 X, Y, Z 来测度，且这 3 个量按前面表中第一行变换。而磁场强度是一个轴性矢量的说法，则表示它的 3 个分量按表中最后一行变化。这里，我把这些分量的因次问题放在一边，否则就钻到物理学的细节上去了。

除了"矢量"这个词以外，哈密顿还引入了在今天物理学中同样起着重要作用的"标量"这个词。标量是在所有坐标变换下不变的简单量，即在坐标改变的情况下，本身或一点不变，或只乘上一个因子。深入下去，还可以区分标量概念中的细微差别。例如，首先考虑空间块或四面体的体积

$$T=\frac{1}{6}\begin{vmatrix} X_1 & Y_1 & Z_1 & 1 \\ X_2 & Y_2 & Z_2 & 1 \\ X_3 & Y_3 & Z_3 & 1 \\ X_4 & Y_4 & Z_4 & 1 \end{vmatrix}。$$

通过计算很容易验明其变换如下：

变　换	平　移	旋　转	反　演	长度单位改变
变　成	T	T	$-T$	$\lambda^3 T$

这种在平移和旋转下不改变而在反射下改变的量，称为第二类

标量，第一类标量则在反演时也不改变。在这个定义中，对第四列给出的因次不加考虑。

很容易建立第一类标量，最简单的例子是 $X^2+Y^2+Z^2$（其中 X, Y, Z 是自由矢量的坐标），以及 $l^2+m^2+n^2$（其中 l, m, n 是自由平面度量的坐标）。事实上，从第一个表及对旋转变换系数考虑方程(3)可以推出这些量在所有运动和反射（不改变长度单位）下保持不变。因此，它们必然具有纯粹的几何意义。我们确已知道，它们代表矢量长度的平方，或在相应情况下代表平面片的面积。

现在要问：如何从给定基本几何图形（矢量和两类标量）的组合中获得同类的其他图形？首先考虑一个十分简单的例子。设 T 是一个第二类标量，例如四面体体积，X, Y, Z 是一个极性矢量坐标。我们考虑 3 个量 $T\cdot X$, $T\cdot Y$, $T\cdot Z$。它们在运动下的变换情况和分量 X, Y, Z 本身一样。但在反演时，由于两个因子都改变了符号，所以它们保持不变。这 3 个量代表轴性矢量。类似地，从轴性矢量 l, m, n，可得到极性矢量 $T\cdot l$, $T\cdot m$, $T\cdot n$。

现在取两个极性矢量 X_1, Y_1, Z_1 和 X_2, Y_2, Z_2，从一个纯粹的解析步骤出发，用以形成各种类型的典型组合。我们将考虑在坐标变换下这些新构成的量的性质，并由此确定它们代表何种几何量。

(1) 我们从和 X_1+X_2, Y_1+Y_2, Z_1+Z_2 出发。显然，它们的变化方式和矢量的分量完全一样，因此，它们代表一个新的极性矢量，它与给定两矢量形成一个与坐标系无关的纯粹几何关系。

(2) 两矢量分量的双线性组合

$$X_1X_2+Y_1Y_2+Z_1Z_2,$$

很容易通过计算验明，它在所有运动与反射下都保持不变，因此是第一类标量，必可给出纯粹的几何定义。

(3) 由分量组成的矩阵

$$\begin{vmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \end{vmatrix}$$

的 3 个子行列式，不难证明，正好和一个自由平面度量或一个轴性矢量的坐标的性质一样，因而必然与给定矢量相联系，而与坐标无关。

(4) 最后考虑 3 个极性矢量及由其 9 个分量形成的行列式

$$\begin{vmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \\ X_3 & Y_3 & Z_3 \end{vmatrix}。$$

在所有运动下它保持不变，但反射时改变符号，所以它定义了一个第二类标量。

我将对这些图形作出几何解释。一旦说明结果后，只要从适当确定的坐标系出发，是不难给出证明的。

对上述(1)的解释。这里定义的两个矢量和是众所周知的。如果两矢量从同一点画出，则从这一点画出的由两矢量组成的平行四边形的对角线代表和(力的平行四边形法则，图 13.3)。

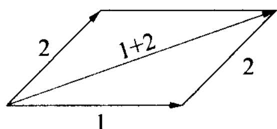


图 13.3

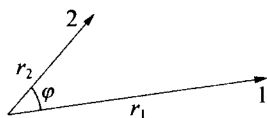


图 13.4

对上述(2)的解释。如果矢量的长度为 r_1 和 r_2，且其方向间的夹角为 φ(图 13.4)，则双线性组合是 $r_1 r_2 \cos\varphi$。

对上述(3)的解释。我们还是考虑一个平行四边形，它的边平行于矢量 1 和矢量 2。并设想它依次按 1 和 2 的方向环行(图 13.5)，这样就得到一个完全确定的自由平面度量，即上述由平行四边形的 3 个坐标确定的自由平面度量，其面积的绝对值由 $r_1 \cdot r_2 |\sin\varphi|$ 给出。

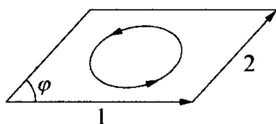


图 13.5

对上述(4)的解释。如果 3 个矢量从一点出发，形成一个平行六面体的 3 条边（图 13.6），其带有适当确定符号的体积，则等于由行列式给出的第二类标量。

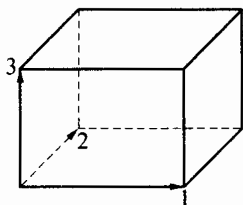


图 13.6

现在讲一下其他文献中处理这些问题的方法。在其他文献里，通常不像我们这样，把在坐标变换下某些解析表达式的性质的研究，即不变量的合理而简单的理论放到第一位。在通常处理中，某些力学和物理学术语都是由格拉斯曼和哈密顿的术语演变而来。通常都说所谓矢量代数、矢量分析，把从给定矢量形成新矢量和标量的法则与普通数量的初等运算法则作比较。

首先指出，正如已经指出的，在第(1)点中出现的运算，可以干脆称为两矢量 1 与矢量 2 的相加。这样命名的合理性，在于符合普通数相加的某些形式规律，特别是交换律与结合律。按交换律，和的定义与所用两矢量 1 和矢量 2 的顺序无关。按结合律，矢量 1、矢量 2 之和同矢量 3 相加，其结果与矢量 1 同矢量 2、矢量 3 之和相加无异。第(2)点和第(3)点所确定的运算称之为相乘，则带有一定的自由性，我们分别称之为内积或标积(第(2)点)与外积或矢积(第(3)点)。事实上，由方程

$$a_1(a_2+a_3)=a_1a_2+a_1a_3$$

表示的所谓乘法对加法的分配律这一重要性质，对两者均成立。例如对内积，我们有

$$\begin{aligned}&X_1(X_2+X_3)+Y_1(Y_2+Y_3)+Z_1(Z_1+Z_3)\\&=(X_1X_2+Y_1Y_2+Z_1Z_2)+(X_1X_3+Y_1Y_3+Z_1Z_3)。\end{aligned}$$

对外积的这个性质可同样简单地导出。对去年冬天我在讲座中曾加以充分讨论的其他的乘积的形式规律①，我可以说乘法交换律

① 见第一卷第一部分第二节。

($a \cdot b = b \cdot a$) 对内积成立而对外积则不成立，因为当矢量 1 和矢量 2 交换时，确定外积的矩阵的子行列式改变了符号。

我可以补充一点，两极性矢量的外积常被简单地定义为一个矢量，而不过分强调其轴的特性。当然，在前面所述一般关系的基础上，可以用一个矢量代替自由平面度量，并得到下面法则：矢量 1 和矢量 2 的外积是一个长为 $r_1 r_2 |\sin\phi|$ 的矢量 3，它垂直于矢量 1 和矢量 2 所在的平面，其正向使矢量 1，2，3 的相互位置对应于正 X，Y，Z 轴的相互位置(图 13.7)。但不要忘记，这个定义本质上依赖于坐标系及长度单位。

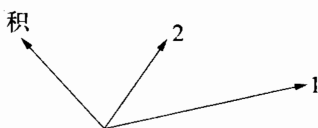


图 13.7

为什么这样牢牢地使用这种矢量分析的语言，我不能充分理解，也许是因为许多人觉得与历史悠久的普通计算运算有这种形式上的雷同是很大的快乐。无论如何，这些矢量运算的名称因普遍容忍而一直使用至今。但是，在选择这些运算，特别是各种乘法运算的确定符号上，已产生了很大的意见分歧。在前面的讲座中①，我已经解释过，尽管做了一切努力，仍然存在着巨大的不一致。在最近的罗马数学大会上已成立一个国际委员会，要求它提出统一的符号。至于委员会成员之间是否能取得任何一致，大量的数学家是否能接受委员会的建议，只能由时间来证明。舒舒服服专门躺在自己那一套上的人们，除了在立法或物质利益的强制力下，是很难调和其分歧观点的。我在这里宁可不谈矢量分析的符号，否则我可能无意地又创造了一个新符号。

我想结束这个讨论而应强调指出，就我们的一般立场而言，普通矢量分析问题仅仅是非常丰富的更一般问题中的一章。例如，滑动矢量、受约束的平面度量、螺旋和力系等，严格地说，都没有在矢量分

① 见第一卷第一部分第四章第三节。

析中考虑。但是，为了真正理解矢量代数本身的运算，实际需要采取较广泛的观点来看待。只有这样，矢量代数运算中所固有的原理，即根据直角坐标各种变换下的性质而确定几何量的原理，才能得到充分的表达。至于涉及这一切问题的文献，我首先要提到我的一篇文章①，其中解释了一般的分类原理，并曾特别将这个原理应用于前曾提及的、罗伯特·波尔爵士的螺旋理论。我还要提到 E·蒂默丁(E. Timerding)和 M·阿伯拉罕(M. Abraham)写的刊于《百科全书》中的文章(分别为："Geometrische Grundlegung der Mechanik eines starren körpers"，《百科全书》第 4 卷第 2 分册，以及"Geometrische Grund begriffe der Mechanik deformierbarer körper"，《百科全书》第 4 卷第 14 分册)。

[**本版附记：**罗马建立的矢量符号统一委员会正如预料没有取得丝毫成功。在随后的剑桥大会(1912 年)上，委员会成员不得不对没有完成任务作出解释，并要求将时间延长到下一次大会。那一次大会本应于 1916 年在斯德哥尔摩召开，但因战争而取消。单位及符号委员会也遭受了同样的命运。它在 1921 年公布了矢量建议符号，立刻招致各方面的最强烈的反对。这个建议分布在 *Zeitschrift für angewandte Mathematik und Mechanik* 第 1 卷(1921 年)第 421 页及随后部分，反对者的评论公布在同一杂志的第 2 卷(1922 年)。今天常用的矢量运算术语有其历史来源，主要来自哈密顿四元数运算及格拉斯曼延拓理论。格拉斯曼的研究难以读懂，一直不为德国物理学家所知，长期以来是少数数学家圈子里的某种深奥的学说。另一方面，哈密顿的思想主要通过麦克斯韦(J. Maxwell)进入美国物理学。但是麦克斯韦在《电和磁的论文》(两卷本，剑桥，1873 年)中，几乎没有例外地在矢量方程中用分量表示法。由于怕别人看不懂，

① *Zeitschrift für Mathematik und Physik*，第 47 卷第 237 页及后面部分，以及《数学年刊》，第 67 卷，第 419 页。收入 F·克莱因 *Gesammelte Mathematische Abhandlungen*，第 1 卷，第 503 页及随后部分。

他几乎不用某种特殊的符号，尽管他认为，为了许多物理思考目的，最好避免引入坐标，应立刻把注意力引向空间点而不是它的 3 个坐标，引向力的方向及量而不是力的 3 个分量。今天物理学家的所谓矢量运算，来自电报工程师赫维赛特(Heaviside)及美国学者 J·W·吉布斯的工作。后者于 1881 年出版了《矢量分析初步》一书。尽管赫维赛特和吉布斯一开始是追随哈密顿的，但他们两人都把格拉斯曼的思想接了过去，用于他们的运算。矢量运算和格拉斯曼的延拓理论，以及哈密顿的四元数运算，是通过这两个作者的著作间接地进入德国物理学的。把矢量运算引入德国物理学家圈子并仿照赫维赛特写法的第一本书，是 A·福普尔(A. Föppl)的 *Einführung in die Maxwell'sche Theorie*，出版于 1894 年。格拉斯曼和哈密顿两人共同之点在于各自的目标都是以有向量本身进行运算，以后才过渡到分量。突出的是两人都把"积"这个词的意义一般化了。这可能是由于他们一开始就把他们的研究同两项以上复数的理论联系了起来(见本书第一卷中关于四元数的表述)。但是两人用的术语完全不同，前面已经指出，术语"线段"、"平面段"、"平面量"、"内积和外积"都出自格拉斯曼，而"纯量"、"矢量"、"纯量积"、"矢量积"出自哈密顿。格拉斯曼的弟子虽然极其正统，但也用其他一些说法代替了其导师的一部分适当的说法。现存的术语是混杂或经过修改的，用以表示单独运算的符号的使用，一直带有极大的任意性。因此，即使对于专家，在数学上如此简单的这个领域，也极度缺乏明确的术语。

本卷"平面上的格拉斯曼原理"一章中所讲的原理，是扫清这种混乱的指导。据此原理，我们可以将格拉斯曼和哈密顿的理论概括如下：在《延拓理论》中，格拉斯曼研究了属于仿射变换群[①]，且使坐标的原点保持不变的不变式理论。他这种研究是建立在他以后发表的 *Vollständige Ausdehnungslehre* 所说的旋转群的基础上的，哈密

---

① 这些变换将在本卷第五部分中讨论。

顿的《四元数》一书也是如此。哈密顿在这里的分析步骤是完全朴素自然的。他没有想到过在正交群的选择上有任何任意性。如果一方面允许反演,即一切坐标轴在原点的反射,另一方面又以多余而予以排斥,那么正如已经解释过的,还会产生其他的分歧。这一切混乱情况最好用“外积(自由平面量)”、“矢量积”及“矢量”的概念来澄清。如果我们取正交变换群而排斥反演,那么就没有划清这3种量的区别。为此,格拉斯曼在 *Vollständige Ausdehnugslehre* 一书中用矢量来表示自由平面量(旋转指向的平行四边形),他称此矢量为平面量之余,完全对应于物理学家记作矢量积的矢量。但是如果容许反演,那么“平面量”和“矢量积”就要看做是等同的几何图形,而不同于矢量图形。这对应于物理学中极向量及轴向量之间的习惯区别。如果我们现在转到仿射变换群上去,那么就不能再把格拉斯曼的自由平面量和矢量积看做是同样的几何量。]

第十四章

导 出 的 流 形

关于基本的几何图形，我想说的就到此为止，现在来研究由这些图形组合而成的较高级的图形。我将从历史谈起，以便使你们对各个世纪的几何发展获得一些概念。

(A) 直到 18 世纪末，只有点被普遍地用为基本图形，其他基本图形也有出现，但未系统化。作为由点导出来的图形，也考虑了曲线和曲面，以及由不同曲线和曲面组成的图形。让我们简要地考虑一下这些图形有过怎样的变化。

(1) 在初等教材，有时甚至在解析几何引论里，似乎几何的整体也仅限于直线、平面、圆锥曲线和二次曲面。当然，这是非常狭隘的观点。即使古希腊人的知识也部分地超过这个范围，因为其中还包括古希腊人当作“几何轨迹”来看的某些高级曲线。不过这些内容从未进入普通的教材。

(2) 约在 1650 年，费马和笛卡儿(Descartes)开创了解析几何，不妨把当时的前后知识情况作一比较。在 1650 年左右，学者们已把几何曲线与力学曲线区分开来，前者特别包括圆锥曲线，但也包括现在称为代数曲线的某些高级曲线；后者包括某些机械确定的曲线，例如由轮子滚动而产生的摆线。这些曲线大部分属于现在所谓的超越曲线。

(3) 两类曲线都包括在稍后将定义的解析曲线范围之内，其坐标 x，y 能用参数 t 的解析函数，即 t 的幂级数来表示。

(4) 近来常常考虑非解析曲线，其坐标 $x=\varphi(t)$，$y=\psi(t)$ 不能展开成幂级数。例如，由没有导数的连续函数定义的曲线就是这样，

这里隐约存在着更一般的、把解析曲线作为特殊情形的曲线概念。

(5) 最后,通过前面讨论过的集合理论的发展[1],一个在此以前未知的概念,即无穷点集合出现了。这是含无穷多个点的一个总体、一个点的聚集,它可能不严格地形成一条曲线,但仍然由某个规律确定。如果我们希望在具体的直觉里找到某些与点集很好对应的东西,我们可以注视银河,仔细搜索,会发现越来越多的星星。这里,抽象点集理论的真实的无穷性,由近似的数学的无穷性所取代了。

在这本讲义的范围内,遗憾的是没有篇幅来更详细地叙述这些内容,特别是无穷小几何与点集理论,尽管它们同样是几何学中的重要部分[2]。不过它们在许多专门讲义与著作中均有系统的论述,因此我只指出它们在整个几何学领域中的地位,以便留出篇幅,更充分地讨论那些其他地方不常论述的内容。

然而,我还是要解释一下解析几何与综合几何之间的区别,因为在讨论上总要涉及这种区别。按它们的原来意义,综合与解析是描述的不同方法。综合从细节着手,从而建立一般概念,最后达到更一般的概念。解析则相反,从最一般开始,分离出越来越多的细节。正是这种意义上的差别,造成了合成化学与分析化学的名称的不同。类似地,在中学几何里,我们谈几何结构的分析,总是假设已找到所需的三角形,然后把给出的问题分成若干个局部问题。

但在高等数学里,这些词却奇怪地有着完全不同的意义。综合几何为研究图形而研究图形,不求助于公式,而解析几何则前后一贯地利用在适当选择坐标系后能写出来的公式。正确的理解是,按着重于图形或公式,这两类几何学只存在着程度的差别。解析几何如完全不讲几何图形,很难被称为几何学;综合几何除非用到适当的公式语言以使其结果得到精确的表达,也很难称得上是几何学。本书中,所用的处理方法表示我们已经认识到这点,因为我们一开始用公

---

① 见第一卷附录第二节。

② 本卷第三部分将包含有关的某些内容。

式,然后询问它们的几何意义。

但在数学中和在其他领域中一样,人们喜欢分派别,因而出现了纯粹综合学派与纯粹分析学派,他们主要强调绝对的“方法的纯粹性”,致使十分片面,违背了研究对象的性质。因此,解析几何常常脱离几何图形,陷于盲目的计算。另一方面,综合几何又人为地避免一切公式,因此,除了发展其本身的一套古怪的语言公式以区别于普通公式外,最终并未带来什么。这种片面夸大若干基本原则以成其学派的倾向,造成了某种僵化的情况。发生这种情况时,各派学科的进步,主要来自“外界人士”的刺激。以几何学来说,首先弄清楚解析与非解析曲线的差别的,就是函数论的研究者,但这种差别却从未受到以上两个学派中任一派代表人物或教科书的足够重视。类似地,正如我们已经看到的那样,传播矢量分析的是物理学家,虽然基本概念可以在格拉斯曼的著作里找到。即使在今天的几何教材中,矢量仍很少作为独立的概念加以阐述。

不断有人建议,应该把几何当作独立的教学课程从数学中分离出来,并为了教学的方便,一般地说应该把数学分为几个科目。事实上,特别是在国外的大学里,已有几何教授、代数教授、微积分教授等等之分。从以上的讨论中,我想得出推断:那样细分是不可取的。相反,有共同研究对象的各个科学分支,应容许尽可能充分地相互交叉。有关的每一个分支,原则上都应看做是数学整体的代表。基于同样的想法,我赞成在数学家和各不同科学学科的代表之间进行最积极的交流。

结束离题的话,继续谈历史发展的情况。

(B) 现在考虑从 1800 年开始出现所谓新几何学而使几何研究受到巨大冲击。今天,新几何学不如称为射影几何,因为后面将充分讨论的射影运算起到了主要作用。新几何学这个名称现在还大量使用着,但实际上是不恰当的,因为从那时起,已经出现了许多更新的发展。射影几何的第一个开拓者,我认为是 J · V · 彭赛列

(Poncelet),他于 1822 年发表了《图形的射影性质》一书。

在射影几何进一步发展之初,也有综合与分析方向之别。第一类代表,我要提到德国人 J · 斯坦纳(J. Steiner)和冯 · 斯坦特(Von Staudt),第二类代表除莫比乌斯外,首推普吕克。甚至今天仍有积极影响的这些人的奠基作是:斯坦纳的 *Systematische Entwickelung der Abhängigkeit geometrischer Gestalten von einader*,冯 · 斯坦特的 *Geometrische der Lage*,莫比乌斯的 *Baryzentrische Kalkül*,普吕克的 *Analytischgeometrische Eutwikelungen*。

如果要我提出这些"新"几何学的最重要的指导原则,我想首先介绍下列思想。

(1) 我要挑选出来讲的彭赛列的主要成就,是他第一次突出了这样一个思想:存在着与点具有同等合理性的其他图形。特别是,在平面上可以选无限直线与点对照,在空间上可以将无限平面与点对照。在大量几何定理中,我们可用"线"字或"面"字代替"点"字。这就是对偶原理。

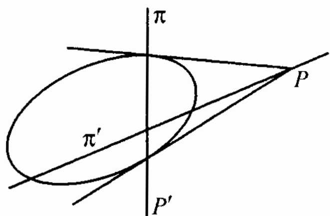


图 14.1

彭赛列的推导从互反的极点及极线理论或圆锥曲线极性理论出发。众所周知,每一个点 P,相对于给定的圆锥曲线,确定一条称为该点极线的直线 π,它可以定义为连接从 P 到圆锥曲线的两条切线之切点的直线(图 14.1)。反之,每条线 π 对应一个极点。从互反关系可得到:如果在 π 上有一点 P',则 P' 的极线 π' 必通过 P。从由圆锥曲线建立起来的、平面上点与线之间的这个一一对应关系,加上类似地由空间二次曲面建立起来的空间点与平面的对应关系,彭赛列得出结论说,可用这种方法将所有涉及位置性质的定理对偶化,而使点与线或点与平面相对应。一个著名的例子是可把关于内接圆锥曲线的六边形的帕斯卡(Pascal)定理,对

偶成关于圆锥曲线外切六边形的布赖恩昌(Brianchon)的定理。

(2) 对偶性原理的深入研究结果,使它逐渐从极性理论中分离出来,并被承认为射影几何的一个特殊组成部分中的推论。这样美妙的系统化,首先出现在格根(Gergonne)和斯坦纳的著作中。

因为以后经常谈到这个学科,我现在先把它简单介绍一下。对偶原理可叙述如下:在几何学的基本概念与基本定理(公理)中,空间里的点和平面,或限于平面上的点和直线总是对称的,即公理以及由公理通过逻辑推导出来的定理,总是成对地出现的。如距离、角等等所谓初等几何的量的关系,起初并不出现在这个学科中。后面我们会看到,这些内容是怎样补进去的。这个结构的组成如下:

(a) 3 种图形:点、(无限)直线、(无限)平面,是最简单的基本图形。

(b) 在这些基本图形之间存在着下列关系(称之为连接定理或公理):两点决定一直线;三不共线的点决定一平面;两平面决定一条直线;三不共线的平面决定一点。巧妙地引入若干额外的(无穷远)元素,可使这些公理无条件地成立。后面将介绍引入这种元素的方法。

(c) 现在构造线性基本图形(即由线性方程解析地给出的图形)。

Ⅰ. 第一类基本图形,各由∞^1 个元素组成。

(α) 在直线上的所有点:直线点域;

(β) 过一直线的所有平面:平面束;

(γ) 在平面上通过一点的所有直线:(平面)线束。

Ⅱ. 第二类基本图形,各由∞^2 个元素组成。

(α) 平面作为它的点的轨迹:点场;

(α′) 平面作为它的线的轨迹:线场;

(β) 通过一固定点的平面:平面束;

(β′) 通过一固定点的直线:直线束。

Ⅲ. 第三类基本图形，各由∞^3个元素组成。

(α) 空间作为它的点的轨迹：点的空间；

(β) 空间作为它的平面的轨迹：平面的空间。

在这个结构里，到处出现完全的对偶性。如果利用给出的基本元素，我们一方面从点出发；另一方面从线(如果关心的是平面几何)或从平面(如果考虑的是空间几何)出发，则可以用两个相互对偶的方法建立起整个射影几何结构。

(3) 如果遵循解析方法，为此而首先问对偶原理在普吕克理论中的出现形式，那么可以用另一种方便的方法把这个结论表示出来。如果常数项不是零，可将平面中的直线方程写成

$$ux+vy+1=0。$$

知道了系数u，v的值，也就确定了这条直线，这些值与流动坐标x，y以对称的形式出现。普吕克的想法是，把u，v当作“直线的坐标”，当作与点坐标x，y是同样合理的，而且有时代替x，y当作变数。从这个新观点出发，x，y具有固定值，该方程则表达一条变直线通过固定点的条件：这个点在直线坐标下的方程。最后，在所用表达式的形式下，没有必要指明它表示哪一个图形。这些量中哪一对当作常量，哪一对当作变量，可以完全不确定。于是这个方程所表达的是点与线的“联合位置”的条件。对偶原理即为：每一个方程中，以x，y为一方，以u，v为另一方，是完全对称的。上面关于连接性公理所固有的对偶性等等，都依赖于这个性质。

在空间，直线方程当然由平面方程

$$ux+vy+wz+1=0$$

所代替。

根据这些讨论，把x，y，z或u，v，w当作基本变量，并相应地把点和平面两个词互换，就可解析地发展几何。这样也就出现了大家所熟悉的几何学的两个方向的发展，很多教材中对这一点都有所

强调,同一页上会并列出现对偶的定理,中间用垂直线分开。请稍微看一下用这个方法导出的高级图形,总是以对偶形式出现,在某种意义上,这些就是上面对偶线性图形的进一步发展。

首先把 x, y, z 当作参数 t 的、确定的、非常数的函数 ϕ, χ, ψ。于是,这 3 个函数代表一条空间曲线,特殊情形下(当 ϕ, χ, ψ 恒等地满足一个常系数线性方程时)可以是一条平面曲线,或甚至(当它们满足两个这种线性方程时)退化成一条直线。同样,考虑 u, v, w 为 t 的函数,我们得到一个单重无穷的平面族,可以十分方便地把它当作可展曲面的切平面。作为特殊情形,以第一种情形来说,所有平面通过一点,即形成一个锥面的包络。以第二种情形来说,所有平面都经过一条固定的直线。

其次,如果把 x, y, z 当作两个参数 t 和 t' 的函数,则得到一个曲面,在特殊情形下,可以退化成一个平面。它的对偶是包围一个曲面的双重无穷个平面,它可能退化成过一点的平面束。

兹将结果列表如下:

$\left.\begin{array}{l} x=\phi(t) \\ y=\chi(t) \\ z=\psi(t) \end{array}\right\}$ 曲　线(平面曲线)(直线)	$\left.\begin{array}{l} u=\phi(t) \\ v=\chi(t) \\ w=\psi(t) \end{array}\right\}$ 可展曲面(圆锥面)(直线)
$\left.\begin{array}{l} x=\phi(t, t') \\ y=\chi(t, t') \\ z=\psi(t, t') \end{array}\right\}$ 曲　面(平面)	$\left.\begin{array}{l} u=\phi(t, t') \\ v=\chi(t, t') \\ w=\psi(t, t') \end{array}\right\}$ 曲　面(点)

这足以作为对偶方式的一个例子。

(4) 甚至在普吕克的研究里,这整个课题的内容,也得到了重大的扩充。正如他把平面方程的 3 个系数当作可变的平面坐标一样,他也曾设想把任何几何图形一般所依赖的若干常量(如二次曲面方

程的 9 个系数)当作这个图形的可变坐标,并研究了可变坐标之间的方程。当然,不能再谈真正意义上的"对偶性",因为对偶性依赖于平面方程和直线方程中系数与坐标对称出现的这个特殊性质。

普吕克自己用上述思想特别研究了空间中的直线,一条空间直线以点坐标的两个方程给出,普吕克将之写成形式

$$x = rz + \rho,\ y = sz + \sigma。$$

把这些方程中的 4 个参量 r, s, ρ, σ 称为空间的直线坐标。不难指明,它们与前面根据格拉斯曼原理从直线上的两点导出的 6 个比 $X : Y : \cdots : N$ 是相关的。普吕克现在考虑这 4 个坐标之间的一个方程 $f(r, s, \rho, \sigma) = 0$,它从空间的四重无穷个直线中分离出三重无穷个直线,普吕克称之为直线丛。我们以前已谈过这个流形中的最简单情况:线性丛。两个方程

$$f(r, s, \rho, \sigma) = 0,\ g(r, s, \rho, \sigma) = 0$$

确定了也被称为射线系的一个线汇。后一名称表示我们所关心的,是两个线丛 $f = 0$, $g = 0$ 所共有的那些直线。最后,3 个同类方程 $f = g = h = 0$ 决定了一个单无穷个直线族,这个直线族覆盖了某个直线曲面。

普吕克在 1868—1869 年名为 *Neue Geometrie des Raumes gegründet auf die Betrachtung der geraden Linie als Raumelement*① 的书中作了这样的表述。该书第一部分快印完的时候他去世了,我作为他的助手,荣获编辑第二部分的机会。

普吕克把任何图形当作一个空间元素,并把它的常数当作坐标的一般原则,后来又引起其他一些有意义的发展。例如,曾在莱比锡工作多年的著名挪威数学家索费斯·李(Sophus Lie),在他的球面几何方面取得了巨大的成就。他把空间元素看做是像直线一样依赖

① 第一及第二部分,莱比锡,1868 年及 1869 年。

于 4 个参数的球。其次，我要提到后一时期斯特迪的 *Geometrie der Dynamem*[1]，这种性质的一系列有趣的研究都和我们已经讨论过的“动力坐标”有关。

(C) 刚才讨论的“新几何学”，主要着重把无限直线和无限平面当作空间元素来对待。但是，格拉斯曼自 1844 年起进行的研究，又超过了这个范围。他突出了有限直线段、平面片、空间块，并按他的行列式原则对它们指定了分量。这一次，我们已经彻底讨论过了。其精彩之处在于，它很适应力学与物理学的需要，对这两个学科的作用远远超过直线几何与对偶原理(这是举例来说)。

当然，在讨论每一个分支的时候，为了使你们得到一个更明确的印象，我把以上这些不同方向的发展分开来讲了，其实它们绝不是彼此有明确的界限。事实是：普吕克更着重于无限直线，而格拉斯曼更着重于线段，但其他图形有时也在他们的研究范围内出现。

不过，我必须强调，格拉斯曼绝不是只限于研究可以直接应用的方面。他自由发挥他的创造直觉，使他的研究范围远远超过这个界限。他的最重要贡献是：他引出 n 个点坐标 $x_1, x_2, \cdots, x_n$ 的一般概念，取代了 3 个坐标 x, y, z，因而成了 n 维空间 R_n 几何学的真正创造者。他遵循他的一般原则，考虑了这样一个较高维的空间，考虑了 $2, 3, \cdots, n+1$ 点坐标的矩阵及其余子式，给出了 R_n 中对应于线段、平面片的一系列基本图形。我曾经说过，格拉斯曼称这样创造出来的一个抽象学科为扩张理论。

现在，R_n 的概念已推广到无穷多个坐标 $x_1, x_2, \cdots$，所以已经有无限维 R_∞ 空间之说了。大家不妨想一下无穷级数的运算：一个幂级数被它的无穷多个系数总体所确定，就这个范围而言，幂级数可作为 R_∞ 中一个点来表示。由此可知，无限维 R_∞ 空间这个概念并不是言之无物。这里值得注意的是，从几何上去谈 n 个甚至无穷个变

---

[1] 莱比锡，1903 年。

量,这种方法已被证明是有实用意义的。有了这种说法,讨论起来就要比抽象的解析表达法生动得多。在这个现象的后面有几分真理?人们思维的发展一般限于二维或三维空间中的经验,新的概念是否是思维的自然产物?这是要由心理学家和哲学家来判定的问题。

言归正传,我想稍微仔细讨论一下高维流形,它作为格拉斯曼基本流行的组合,特别是矢量的组合,可以和我们曾讨论过的点、平面等等的组合并列。这里,我们进入了实矢量分析的进一步研究。由于哈密顿的功劳,这已成为力学和物理学中最有用的工具之一。我向你们推荐哈密顿的“四元数基础”,以及前已提及的杰出的美国数学家吉布斯的“矢量分析”①。

除了我们已熟悉的矢量与标量概念外,这里还要加一个新概念,就是这些量与空间的点的联系。对空间的每一个点,我们指定一个确定的标量 $S = f(x, y, z)$,于是我们说形成一个标量场。另一方面,对空间每一个点,我们附以一个确定的矢量

$$X = \phi(x, y, z), Y = \psi(x, y, z), Z = \chi(x, y, z),$$

并称这些矢量的总体为一个矢量场。

以这种方法,引出了现代物理学中处处都用到的两个最重要的几何概念。其广泛应用的例子很多,只要回忆几个就够了。始终被当作位置函数看的质量分布密度、温度和一个连续延伸系统的势能,都是标量场的例子。在每一点有一个确定的力作用的力场,是矢量场的典型例子。我还可以补充下列例子。在弹性理论中,当我们对变形体的每一点指定一个表明该处位移的大小与方向的有向线段时,就有一个变形体的位移场,它也是一个矢量场。类似地,流体动力学中的速度场,电动力学中对每一点指定一个确定的电矢量与磁矢量的电磁场,都是矢量场的例子。因为在电磁场中每一个点上都

① E·B·威尔逊(E. B. Wilson)编辑,纽约,1901年。

可以把轴性的磁场强度矢量和电场强度的极性矢量结合而成一个螺旋,所以电磁场也可以作为一个螺旋场的例子。

哈密顿曾说明,怎样以最简单的方式把微积分学方法应用于这些场中。为此,关键是要说明,空间一点的位移方向由微分 dx, dy, dz 的比例确定,微分 dx, dy, dz 代表一个自由矢量,即它们在坐标变换下的性质和自由矢量的分量一样。它们是通过点 x, y, z 的一个小线段的坐标过渡到极限而产生的,由此不难推出上述道理。

更为重要但也较难掌握的是第二点,即偏微分符号

$$\frac{\partial}{\partial x},\ \frac{\partial}{\partial y},\ \frac{\partial}{\partial z}$$

也具有自由矢量的分量的特性,即如转到一个新的直角坐标系 x', y', z',新符号$\frac{\partial}{\partial x'}$, $\frac{\partial}{\partial y'}$, $\frac{\partial}{\partial z'}$相对于原符号的性质与一个矢量(特别是一个极性矢量)的坐标变换性质一样。

如果用坐标系的旋转

$$\begin{cases} x' = a_1 x + b_1 y + c_1 z, \\ y' = a_2 x + b_2 y + c_2 z, \\ z' = a_3 x + b_3 y + c_3 z \end{cases} \tag{1}$$

就马上明白了。以前已指出过,这些旋转公式具有这样的特性,只要对其中系数进行行与列的交换,即可求得它们的解

$$\begin{cases} x = a_1 x' + a_2 y' + a_3 z', \\ y = b_1 x' + b_2 y' + b_3 z', \\ z = c_1 x' + c_2 y' + c_3 z'。 \end{cases} \tag{2}$$

如果我们有 x, y, z 的任一函数,通过(2)式,我们可将其表示为 x', y', z'的函数。据已知的微分法则,我们有

$$\frac{\partial}{\partial x'}=\frac{\partial}{\partial x}\cdot\frac{\partial x}{\partial x'}+\frac{\partial}{\partial y}\cdot\frac{\partial y}{\partial x'}+\frac{\partial}{\partial z}\cdot\frac{\partial z}{\partial x'},$$

$$\frac{\partial}{\partial y'}=\frac{\partial}{\partial x}\cdot\frac{\partial x}{\partial y'}+\frac{\partial}{\partial y}\cdot\frac{\partial y}{\partial y'}+\frac{\partial}{\partial z}\cdot\frac{\partial z}{\partial y'},$$

$$\frac{\partial}{\partial z'}=\frac{\partial}{\partial x}\cdot\frac{\partial x}{\partial z'}+\frac{\partial}{\partial y}\cdot\frac{\partial y}{\partial z'}+\frac{\partial}{\partial z}\cdot\frac{\partial z}{\partial z'}。$$

从(2)式立即可得 x, y, z 对 x', y', z' 的导数,于是有

$$\frac{\partial}{\partial x'}=a_1\frac{\partial}{\partial x}+b_1\frac{\partial}{\partial y}+c_1\frac{\partial}{\partial z},$$

$$\frac{\partial}{\partial y'}=a_2\frac{\partial}{\partial x}+b_2\frac{\partial}{\partial y}+c_2\frac{\partial}{\partial z},$$

$$\frac{\partial}{\partial z'}=a_3\frac{\partial}{\partial x}+b_3\frac{\partial}{\partial y}+c_3\frac{\partial}{\partial z}。$$

和(1)式相比较,表明它与点坐标变换公式一致,因此与矢量分量的变换公式一致。

还有一个简单得多的计算也可以证明,在坐标系平移时,3 个符号 $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$, $\frac{\partial}{\partial z}$ 不改变,但在反演变换下改变符号,于是上述性质得证。说实在的,我们没有考虑长度单位改变时的情况,没有考虑因次。如果加以计算,我们会发现偏微分符号的因次为 -1,因为坐标的微分在分母中出现。

我们现在用哈密顿矢量算符 $\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$ 进行前边用矢量所进行的那些运算。首先要指出,作用在函数 $f(x, y, z)$ 上的算符 $\frac{\partial}{\partial x}$ 运算的结果,即 $\frac{\partial f}{\partial x}$ 可以用符号形式称为 $\frac{\partial}{\partial x}$ 与 f 的乘积,因为就这里所涉及的乘法的形式规律,特别是分配律

$$\frac{\partial(f+g)}{\partial x}=\frac{\partial f}{\partial x}+\frac{\partial g}{\partial x}$$

而论，对这种乘积是成立的。

现在，设给出一个标量场 $S = f(x, y, z)$，用哈密顿算符的分量按刚才说明的意义乘这个标量，即设我们形成矢量

$$\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}。$$

前面我们已看到，一个标量乘以一个矢量仍是一个矢量。因为在这个定理的证明中只用到我们的算符乘法所保持的那种乘法性质，于是推出，这个标量场的 3 个偏导数确定了一个依赖于 x, y, z 的矢量，因此是一个矢量场。这个标量场与矢量场之间的联系与坐标系的选择无关。将矢量的符号改变后，借用气象学的术语，这个矢量场称为标量场的梯度。因此，在报纸上常见的气象图里，每一点的空气压力用一个标量场 S 表示，画出了 S=常量的曲线，并标上对应的 S 值。于是，梯度给出了气压下降最快的方向，并始终垂直于这些曲线。

由于矢量分量 X, Y, Z 总能形成一个标量 $X^2+Y^2+Z^2$，因此，从标量的梯度可以得到一个新的标量场

$$\left(\frac{\partial f}{\partial x}\right)^2+\left(\frac{\partial f}{\partial y}\right)^2+\left(\frac{\partial f}{\partial z}\right)^2。$$

它必然以一种和坐标系无关的方式与梯度，从而与原来的标量场相联系。此标量等于梯度的长度的平方，或称为标量场 f 的斜率的平方。

应用同样的原则，使每个分量形式地自乘，即施以两次同样的运算，就可从矢量算符 $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$ 得到一个算符标量，这样产生了运算

$$\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}。$$

因而，它具有标量的特性，即它在坐标变换下是不变量。如果用这个

标量乘以标量场 f，必然又得到一个标量场

$$\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}+\frac{\partial^2 f}{\partial z^2},$$

它与前者的关系是独立于坐标系的。如果设想一个流体在一个场里流动，其初始密度为 1，其在每一点的速度由 f 的梯度给出，则在第一个瞬时 $\mathrm{d}f$，每点密度的增加等于这个标量乘以 $\mathrm{d}t$。因此，我们称

$$-\left(\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}+\frac{\partial^2 f}{\partial z^2}\right)$$

为 f 的梯度的散度。

以前，按拉梅(Lamé)的习惯，常称一个标量场 $S=f(x,\ y,\ z)$ 为一个点函数，而称与它相联系的第一个标量场 $\left(\frac{\partial f}{\partial x}\right)^2+\left(\frac{\partial f}{\partial y}\right)^2+\left(\frac{\partial f}{\partial z}\right)^2$ 为第一微分参数，称另一标量场 $\left(\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}+\frac{\partial^2 f}{\partial z^2}\right)$ 为第二微分参数。

用类似的方法，现在把我们的矢量算符与一个给定的(极性)矢量场

$$X=\phi(x,\ y,\ z)\qquad Y=\chi(x,\ y,\ z)\qquad Z=\psi(x,\ y,\ z)$$

联系起来。当然，我们会用已建立的两个矢量的两类乘法来实现这一点。

(a) 通过内积得到一个标量，用已熟悉的算符乘法记号，可将其写成

$$\frac{\partial X}{\partial x}+\frac{\partial Y}{\partial y}+\frac{\partial Z}{\partial z}。$$

因为这个结果当然也依赖于 $x,\ y,\ z$，所以它也代表一个标量场，它

和原给的矢量场的关系与坐标系无关。按上面已确定的定义,称为这个矢量场的散度。

(b) 通过外积得到矩阵

$$\begin{pmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ X & Y & Z \end{pmatrix}。$$

它的 3 个行列式为

$$\frac{\partial Z}{\partial y}-\frac{\partial Y}{\partial z},\ \frac{\partial X}{\partial z}-\frac{\partial Z}{\partial x},\ \frac{\partial Y}{\partial x}-\frac{\partial X}{\partial y}。$$

按前面的说明,这些行列式定义了一个平面度量,或在相应情况下定义了一个轴性矢量或一个轴性矢量场。这两个矢量场之间的关系,仍与坐标系的选择无关。按麦克斯韦的定义,这个场称为原来场的旋度。

通过系统的几何研究,我们得到了物理学家在各种矢量场的研究中必须用到的各种量。但我们所研究的是纯粹几何学。我必须再三强调的原因,是因为这些内容常常被认为属于物理学的范围,因而是物理书或物理课上讨论的,而不是几何学的讨论内容。其实这种态度是完全没有理由的,只能当作历史发展的残痕才能理解。时机成熟时,物理学就不得不创造它所需要的,而又不能从数学里随手拿来的工具。

这里也存在着我在分析里常常提及的同样的误解。在时间的推移过程中,物理学提及了各种各样需要数学来解决的问题,因此常常对数学创造宝贵的刺激。但时至今日,数学的教学,特别是中学数学的教学,对这个变化仍未注意。它还是照多少个世纪以来所遵循的老规矩办,让物理学费劲地自己来解决自己的工具问题,尽管这些问题可以为数学教学提供比传统课题要合适得多的材料。你们看到,在知识界也存在惯性定律,一切都沿旧的直线轨道移动,一切改变,一切向新的与现代方法的过

渡，都遇到强大的阻力。

第一部分的主要内容就到此为止，这一部分已对我们讲了各类几何流形、几何对象。下面将讨论一个特殊的几何方法，它对这些流形的更精确的研究具有十分重大的意义。

第五部分

几 何 变 换

现在我们要讲的是科学的几何学中最重要的一个章节。我要在这里特别指出的是，这一章的基本思想和较简单的部分可以为中学教学提供十分有益的材料。可以说，几何变换无非是简单的函数概念的推广，而后者正是现代数学教改运动所力图推行的中心内容。

我的讨论从组成几何变换的最简单一类的点变换开始。点的变换使点保持为空间元素，即变换使每个点变成对应的另外的点，而其他类变换则把点变成直线、平面、球等等其他空间元素。这里，又要把解析方法放到第一位，因为我们常常能用解析方法对事实作出表述。

点变换的表达式是分析中所谓新变量 x', y', z' 的引入：

$$\begin{cases} x' = \varphi(x, y, z), \\ y' = \chi(x, y, z), \\ z' = \psi(x, y, z)。 \end{cases}$$

我们可以用两种几何方式解释这个方程组，又可以称之为主动的和被动的方式。被动方程组，表示坐标系的一个改变，即对具有给定坐标 x, y, z 的一个点，指定新的坐标 x', y', z'。在讨论直角坐标系的变换时，我们已经了解了这个意思。对一般函数 φ, χ, ψ 而言，这些公式当然包含更多其他类型的坐标系的变换，例如三线坐标、极坐标、椭圆坐标的变换等。

与此相反，主动，解释为保持坐标系固定而改变空间。对每个点 x, y, z，令点 x', y', z' 与之对应，因而事实上是一个空间变换。今后我们将采用这个概念。

如果再考虑以前代表平移、旋转、反射或长度单位改变的 4 组公式(按被动意义，参阅第一部分，第二章第三节)，而现在用主动意义来解释它们，那就得到点变换的第一批例子。不难看到，前两组公式，相对于不可移动的坐标系，分别代表空间(想象为刚体)的平移和绕点 O 的旋转。第三组公式，给出空间点对原点 O 的反演(相对于 O，每点 x, y, z 变为与它对称的点 $-x$, $-y$, $-z$，见图 1)。

最后一组公式表示以 O 为中心的空间的相似变换。

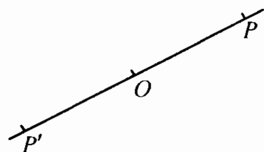


图 1

我们现在从一组特别简单的点变换，即所谓仿射变换开始进行研究，这一组变换把前述一切情形皆作为特例包括在内。

第十五章

仿射变换

当 x', y', z' 是 x, y, z 的线性整函数时,就解析地定义了一个仿射变换

$$\begin{cases} x' = a_1 x + b_1 y + c_1 z + d_1, \\ y' = a_2 x + b_2 y + c_2 z + d_2, \\ z' = a_3 x + b_3 y + c_3 z + d_3。 \end{cases} \tag{1}$$

这个名字起源于莫比乌斯和欧拉,意味着在这样的变换中,无穷远点仍对应于无穷远点。于是,在某种意义下,空间的"末端"保持不变。事实上,这些公式立即表明,x', y', z' 与 x, y, z 一起变为无穷。后面要讨论的一般射影变换则相反,那里 x', y', z' 是分式线性函数,因而某些有限点将变到无穷。仿射变换在物理学中以"均匀变形"的名称起到重要作用。"均匀"(与非均匀相比)意味着系数与所考虑的空间内的位置无关。"变形"告诉我们,任何物体的形状一般通过变换会改变。

变换(1)包括:平行于 3 个坐标轴移动 d_1, d_2, d_3 的平移,加上保持坐标原点不变的齐次线性变换(中心仿射变换)

$$\begin{cases} x' = a_1 x + b_1 y + c_1 z, \\ y' = a_2 x + b_2 y + c_2 z, \\ z' = a_3 x + b_3 y + c_3 z。 \end{cases} \tag{2}$$

这一变换更便于运用。我们从方程组(2)的考虑开始。

(1) 我们考虑解方程组(2)的可能性。正如行列式理论所表明

的，关键在于变换系数组成的行列式

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \tag{3}$$

是否为零。后面我们再考虑 $\Delta=0$ 的情况，现在将假设 $\Delta\neq 0$，于是(2)式具有的形式为

$$\begin{cases} x = a_1'x' + b_1'y' + c_1'z' \\ y = a_2'x' + b_2'y' + c_2'z' \\ z = a_3'x' + b_3'y' + c_3'z' \end{cases} \tag{4}$$

的唯一解，其中 a_1'，…，c_3'是Δ 的代数余子式除以Δ。因此，每点 x，y，z，可以说不仅对应一点 x'，y'，z'，而且只对应这一点，而从 x'，y'，z'到 x，y，z 的变换也是仿射变换。

(2) 我们现在问，空间流形在这些仿射变换下的情况如何？我们先考虑平面

$$Ax + By + Cz + D = 0。$$

将(4)式的 x，y，z 代入其中，作为对应流形的方程，我们得

$$A'x' + B'y' + C'z' + D' = 0,$$

其中 A'，…，D'是变换系数和A，…，D 的某些组合，根据(1)式，我们看到第二个平面上的每一个点从第一个平面上的一个适当点产生。因此，每一个平面对应于另一个平面。因直线是两平面的交线，故必然得出每条直线对应另一条直线。莫比乌斯称这种性质的变换为共线变换，因为它们表达了 3 个点的“共线性”，即位于一条直线的性质。因此，仿射变换是共线性的。

如果用同样的方法研究二次曲面

$$Ax^2 + 2Bxy + Cy^2 + \cdots = 0,$$

将方程组(4)的 x, y, z 代入,我们得到一个二次方程。因此,仿射变换将每个二次曲面变成另一同类曲面。同样地,将 n 次曲面变成另一个同次曲面。

以后,我们会对那些对应于球面的曲面表现出特别的兴趣。首先,它们应是二次曲面,因为球面是这一类中的特殊曲面。但因球面上所有点都是有限的,所以它们当中没有点能变换至无穷远,变换后的曲面必然是整个处于有限区域内的二次曲面,即必须是椭球面。

(3) 我们看看具有分量 $X=x_1-x_2$, $Y=y_1-y_2$, $Z=z_1-z_2$ 的自由矢量将发生怎样的变化。将(2)式用于点 1 和点 2 的坐标,对于相应线段 $1'2'$的分量 $X'=x_1'-x_2'$, $Y=y_1'-y_2'$, $Z'=z_1'-z_2'$,我们有

$$\begin{cases}X'=a_1X+b_1Y+c_1Z,\\ Y'=a_2X+b_2Y+c_2Z,\\ Z'=a_3X+b_3Y+c_3Z。\end{cases}\tag{5}$$

由此可知,这些新分量只依赖于 X, Y, Z,而不依赖于坐标 x_1, y_1, z_1; x_2, y_2, z_2 的个别值,即所有具有同样分量 X, Y, Z 的线段 1 2,对应于具有同样分量 X', Y', Z'的线段 $1'2'$。换句话说,仿射变换将自由矢量变成另一个自由矢量。这个结论的内容比一条直线总对应另一条直线的结论要丰富。设我们在两条平行线上取具有同样正向的两相等线段。由于这些线段代表同样的自由矢量,因此对应的两线段必然表示一个同样的矢量,即必平行、相等并有同样的正向(图 15.1)。每个平行线族仍对应于平行线族,而其上的相等线段,对应于相等线段。这些性质是很重要的,因为不难指出,一般情况下,仿射变换会改变线段的绝对长度和两直线间的交角。

(4) 让我们考虑在同一条直线上不等长的两个矢量。乘以标量后,其中之一可以变为另一个矢量。因为(5)式中的 X', Y', Z'是 X, Y, Z 的齐次线性函数。对应的两矢量将相差同样的一个标量因子,这表明它们之间的长度关系和原来两矢量的一样。我们可以将

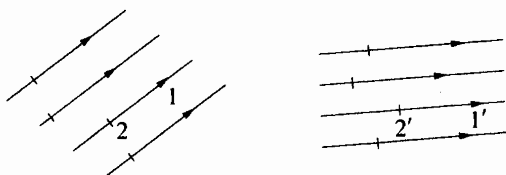


图 15.1

这一点叙述如下：在一个仿射变换下，对应的两条直线是“相似”的，即在两条直线上对应的线段有相同比值。

(5) 最后比较两个四面体的体积

$T=(1,2,3,4)$ 和 $T'=(1',2',3',4')$。我们有

$$6T'=\begin{vmatrix} x_1' & y_1' & z_1' & 1\\ x_2' & y_2' & z_2' & 1\\ x_3' & y_3' & z_3' & 1\\ x_4' & y_4' & z_4' & 1\end{vmatrix}$$

$$=\begin{vmatrix} a_1x_1+b_1y_1+c_1z_1 & a_2x_1+b_2y_1+c_2z_1 & a_3x_1+b_3y_1+c_3z_1 & 1\\ a_1x_2+b_1y_2+c_1z_2 & a_2x_2+b_2y_2+c_2z_2 & a_3x_2+b_3y_2+c_3z_2 & 1\\ a_1x_3+b_1y_3+c_1z_3 & a_2x_3+b_2y_3+c_2z_3 & a_3x_3+b_3y_3+c_3z_3 & 1\\ a_1x_4+b_1y_4+c_1z_4 & a_2x_4+b_2y_4+c_2z_4 & a_3x_4+b_3y_4+c_3z_4 & 1\end{vmatrix},$$

或应用已知的行列式相乘的定理，有

$$6T'=\begin{vmatrix} a_1 & b_1 & c_1 & 0\\ a_2 & b_2 & c_2 & 0\\ a_3 & b_3 & c_3 & 0\\ 0 & 0 & 0 & 1\end{vmatrix}\cdot\begin{vmatrix} x_1 & y_1 & z_1 & 1\\ x_2 & y_2 & z_2 & 1\\ x_3 & y_3 & z_3 & 1\\ x_4 & y_4 & z_4 & 1\end{vmatrix}。$$

第一个因子是 Δ，第二个是 $6T$，故有 $T'=\Delta\cdot T$。在仿射变换下，所有四面体体积，从而所有空间体积(作为四面体体积之和或此和之极限)，变为乘以代换行列式 Δ 这个常数因子。

这几个从仿射变换的解析定义推出来的定理，足以使我们对这种变换获得一个清晰的几何概念。它们的证明比通常给出的要简单，因为我们从矢量概念中已获得适当的表达手段。

如果从坐标 x, y, z 的空间 R 内的球面出发，则可对仿射变换获得一个最清晰的概念。正如我们已经知道，这个球永远对应于坐标 x', y', z' 的空间 R' 内的一个椭球。如果考虑球内的一个平行弦族，则根据第(3)点，它们将对应于椭球的平行弦族(图 15.2)。又因对应的直线是相似的(根据第(4)点)，故球弦的中点必然对应于椭球弦的中点。因为球的弦族的中点在一个平面上，故按照第(2)点的基本性质，椭球弦的各中点必然也在一个平面上，我们称之为椭球的一个直径平面。现在，球的所有直径平面、包含其中心 M，它平分通过它的每个球的弦(球的直径)，因此，对应的 M'(椭球的中心)位于每一个直径平面并平分通过它的每条弦(椭球的直径)。

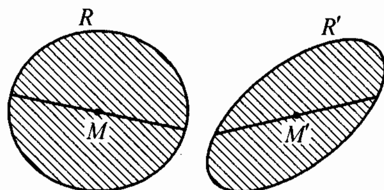


图 15.2

注意，什么对应于球的 3 个互相垂直的直径平面系，也是很重要的。这个平面系显然有着这样的特性，即 3 个平面中每一个平分平行于另两平面交线的弦。这个性质在仿射变换下仍得以保持。因此，球面的无穷多组 3 个互相垂直的直径平面，每一组都对应于椭球的 3 个直径平面，使得平行于其中两平面交线的弦被第 3 个平面平分。这样的平面组称为 3 个共轭直径平面，它们的交线称为 3 条共轭直径。

我假设你们已经知道一个椭球包含有 3 条所谓主轴，即 3 条互相垂直的共轭直径。根据前述，空间 R 内球面必然有 3 条彼此垂直的直径，在仿射变换下对应于它们。为简单起见，设椭球与球的中心分别为 R' 和 R 的原点，通过适当旋转，使这两组互相垂直的直径分

别为空间 R'中的 x', y', z'轴和空间 R 中的 x, y, z 轴。至于这里可任意设想旋转的是空间轴或坐标轴，在任一情况下，这个过程受我们考虑过的特殊类型的线性齐次坐标变换的影响。由于接连几个线性齐次变换总是产生同类型的另一变换，因此将 R 变成 R'的变换方程，在新坐标下，将具有形式(2)：

$$\begin{cases} x' = a_1x + b_1y + c_1z, \\ y' = a_2x + b_2y + c_2z, \\ z' = a_3x + b_3y + c_3z. \end{cases}$$

这样选择新坐标系，就使 x'轴对应于 x 轴，即当 $y = z = 0$ 时，也有 $y' = z' = 0$。由此推出 $a_2 = a_3 = 0$，同理可得 $b_1 = b_3 = c_1 = c_2 = 0$。如果不考虑附带的旋转，则每个仿射变换就是所谓的"纯粹仿射变换"

$$\begin{cases} x' = \lambda x, \\ y' = \mu y, \quad \text{其中 } \Delta \lesseqgtr 0, \\ z' = \gamma z, \end{cases} \tag{6}$$

或按物理学家的说法，是一个纯应变。我们可用下列简单方法对这些方程作出几种解释：空间沿平行于 x 轴方向被拉长 λ 倍(如 $|\lambda| < 1$，则为压缩)，如 $\lambda < 0$，则加上对 x 轴的反射；对其他两个坐标方向，则分别为 μ 倍和 ν 倍。简言之，我们可以把一个纯仿射变换当作空间沿 3 个互相垂直的方向均匀拉伸。这就是最清楚的一个几何解释。

如果可用倾斜的平行坐标轴，这些关系就更为简单，在空间 R 任取一直角或斜角的坐标轴 x, y, z，不改变原点的位置，我们选用对应于它们的线为 R'内的轴 x', y', z'。新轴一般是斜的，具有固定原点的、从直角坐标变换成斜角坐标的变换公式，是(2)式的线性齐次方程。由于两个这样变换的组合仍是同类的变换，因此即使应用上述斜角坐标系，仿射变换的方程必然具有(2)式。但在我们所选

的坐标轴下，它们必然把 R 的 3 个轴变换为 R' 的那 3 个轴。在重复上面的讨论后，可以得出结论：方程组实际上化为(6)式。因此，如果把两个对应的 3 条轴选用各自的(斜)平行坐标系的轴，仿射变换的方程具有简单的特殊(6)式。

与我们的讨论相联系，提出了一个问题，就是要找一个能进行仿射变换的机械装置。这个问题是我在 1908—1909 年冬讲授力学课时提出来的，后来找到了一个漂亮的解决方案。最好的方案是 R·雷马克(R. Remak)的方案，他既考虑了基本思想，又考虑这部机械装置的适当形式。他利用一种所谓"Nürnberg 剪刀串"作为运动装置，这种装置是用构成一串相似平行四边形的杆铰接而成的链。相连两平行四边形所共有的各角点 S_0，S_1，S_2，…，在铰接系统的各种变形下，在连接的直线 g 上形成相似的点列，变成平行四边形的公共对角线(图 15.3)。如果我们把 3 把这样的"剪刀"在任一角点 S 处连接起来形成一个三角形，则所有角点组成的点系将随整个铰接系统的每一变化而发生仿射变换。我们用剪刀的两条对角线作成一个斜角坐标系(图 15.4)，就会明白这一点。如果在三角形的任两点 S 之间插入另外的同类的剪刀，并考虑它们的角点 S，那就能得到经过同样仿射变换的附加的点(图中，这些剪刀由其对角线表示)。根据这个原则，我们能为各种可变的仿射系统建立平面与空间的模型[①]。

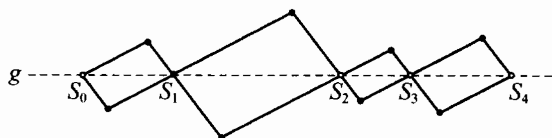
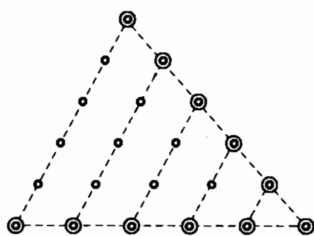


图 15.3

① 一系列这样的模型已在莱比锡马丁·席林(M. Schilling)出版社出现，参阅 F·克莱因及席林夫人著：*Modelle zur Darstellung affiner Transformationen in der Ebene und im Raume*, *Zeitschrift für Mathematik und Physik*，第 58 卷，第 311 页，1910 年。


图 15.4

我们不再进一步讨论仿射变换的性质，而指出如何使用这些变换。

首先，它们为新几何定理的发现提供了一个绝妙的工具。上面解释过的由球到椭球的仿射变换，使我们从球的已知性质推出关于椭球的新定理。例如，如果我们作出球的 3 条互相垂直的直径，加上在这些直径端点上的 6 个切平面，我们得到一个体积为 $J=8r^3$ 的外切立方体，其中 r 为球的半径。仿射变换显然把球的每个切平面变成椭球的切平面，借助上面的定理可推出，空间 R 内的立方体对应于空间 R' 内一个外切于椭球的平行六面体，它的各面在 3 条共轭直径的端点处切于椭球，是平行于对应的直径平面的，而它的棱则平行于那些直径(平面上的圆和椭圆保持着类似的关系，见图 15.5)。此结论的逆显然也成立：每个外切于椭球的平行六面体，按上面说明的方法，对应一个外切于球的立方体，因为椭球的 3 个共轭直径对应于球的 3 条互相垂直的直径。我们已经知道，在仿射变换下，体积要乘以代换的行列式 Δ，故上述外切椭球的平行六面体的体积由公式 $J'=J\cdot\Delta=8r^3\cdot\Delta$ 给出。这个公式显然与平行六面体的位置无关，故不论属于怎样的 3 条共轭直径，上述平行六面体的体积为同一常量。如果我们选择 3 条互相垂直的主轴为所述的共轭直径，则所得长方体之体积为 $8abc$，其中 $2a$，$2b$，$2c$，是主轴之长。我们用这个方法定出了该体积常数，于是我们的定理取如下形式：外切于椭球且其面平行于共轭直径平面的所有平行六面体，有同样的体积 $J'=8abc$，其中 a，b，c 是半轴长。为了证明这

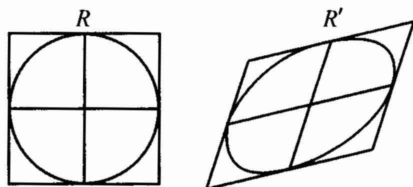


图 15.5

个定理对所有的椭球均有效，只要指明每个椭球可以通过一个球的仿射变换而变成就可以了。这一点可以从仿射变换(6)式的方程立即推得。这些方程表明，椭球的轴的相互比为 $\lambda:\mu:\nu$，这里的 λ，μ，ν 为 3 个任意数。

虽然我们只讲了仿射变换对理论几何的应用，举了一个简单的例子，但我想进一步强调，仿射变换在实践中具有更重大的意义。

首先看看物理学的需要。应该指出，仿射变换在弹性理论、流体力学中起着重大的作用。我不需要作任何夸张，因为每个从事这些学科的人都充分了解，只要一考虑到足够小的空间元素，问题就会化为齐次线性变换。

这里，我想花些时间讨论仿射变换在物理学家和数学家都用得到的正确作图上的应用。就平行投影而论，基本上只涉及空间的仿射变换。遗憾的是，在正确作图这个方面犯有许多错误。在数学教科书的空间图形中，或在物理书的仪器图里，你会找到难以置信的错误。只说一个例子。地球上的赤道常常被画成两个相交的圆弧(图 15.6)。这当然是错误的，正确的画法应如右面所看到的，总是一个椭圆。

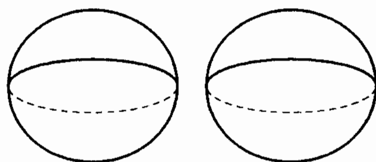


图 15.6

正确作图的原则是：所画的图形应是从一点到图的平面上的投影。如果我们想象中心点在无穷，即如果借助一个平行射线束来作图，关系就简化了。这正是我们感兴趣的情况。顺便讲一点画法几何学。我不准备做系统的讨论，只指出它在整个几何学里的地位，我们也不给出证明的细节。

请先研究平面图形的表示，即通过一个平行线来从一个平面 E 到另一个平面 E' 的投影。为此，把 E 与 E' 的交线选为 x 轴，其上一点为原点 O(图 15.7)。在 E 上任选 y 轴，例如垂直于 x 轴而通过 O，y' 轴则作为 y 轴通过平行线束在 E' 上的投影，因而在 E' 内，一般

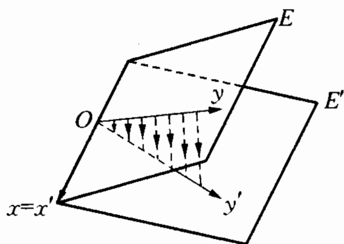


图 15.7

有一个斜角坐标系。于是，E 与 E' 的两对应点的坐标满足关系

$$x' = x,\ y' = \mu \cdot y,$$

其中 μ 是一个依赖于平面与线束所给定位置的常数。因此，这实际上是一个仿射变换。这些方程的证明简单得不需做任何说明。它们是一般(6)式的特例；在这里取 $\lambda = 1$，从而 $x' = x$。当然，这是由于 x 轴是原平面与作图的平面的交线，因而其上每一点的像都与自身重合。如果将原来对空间推导的定理特殊化，用到平面上，例如 E 上的每一个圆对应于 E' 内的椭圆等等，则立即可得图形的所有本质特性。

现在自然会提出相反的问题：如果两平面 E 与 E' 间给出了仿射关系，能使其中一个平面成为另一个平面的平行投影吗？为决定这个问题，设从 E 中任意圆及其在 E' 中对应的椭圆出发(也可用两个对应的椭圆来代替)。圆的中心 M 将对应于椭圆的中心 M' (图 15.8)。如果将 E 内的圆放入平面 E'，使其中心与 M' 重合，则它将与椭圆相交 4 个点或完全不相交。为了简单起见，相切的极限情形不加以考虑。

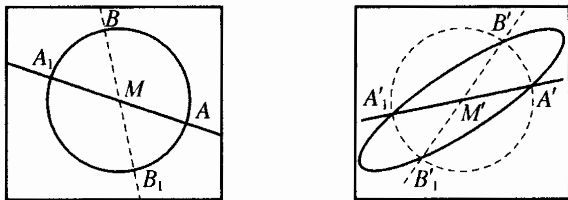


图 15.8

第一种情况如图所示。我们考虑通过 E' 内 4 个交点的椭圆的两条直径 $A'A_1'$，$B'B_1'$。对应于它们的，是在 E 内的圆的两条直径

AA_1，BB_1（按我们的取法，两者相等）。因此，据仿射变换的一般性质，在 AA_1 与 $A'A_1'$ 和 BB_1 与 $B'B_1'$ 上的对应线段是相等的。如果将平面 E 叠合在 E' 上，使 M 与 M' 重合，且使这些线段对之一（例如 AA_1 与 $A'A_1'$）重合，然后把 E 绕这条作为轴的线旋转到空间，即得到两平面间的仿射变换，并使它们交线上的每一点对应于自己。于是，虽然我不给出证明，仍不难看到，无论两平面的夹角是什么，对应点的连线是相互平行的，即两平面间的仿射变换确实通过平行投影而实现。

如果我们的圆与椭圆不相交，即圆的半径小于椭圆的短半轴或大于长半轴，则用分析的话来说，两公共直径就是虚的，不能用于作图，因此无法继续作图。如仍想找出平行投影，则必须应用一个相似变换，把圆放大或缩小，直到第一种情况出现为止。我们在作图中常用相似变换以"改变尺寸"。因此，最终得到主要定理：两平面间的任何仿射关系，可以用无穷多个把相似变换与平行投影组合在一起的方法而得到。

我们现在讨论通过平行投影把空间整个变到平面的问题，这个问题比把一个平面映射到另一个平面的问题要重要和有意义得多。为了避免繁琐，我们总是同意用一个相似变换把图形放大或缩小。这就提出了画法几何中所谓轴测映射方法的问题。这个方法有非常重要的实用价值。如果物体到照相机的距离足够远，则每个像都十分近似于一个轴测映射（严格来说是一个中心投影）。在大多数情况下，当我们希望对一个空间几何图形、物理仪器、建筑构件等等绘图时，严格的轴测映射是特别有用的。在各种轴测映射绘图中，也是在教学中直接有用的一些有趣的例子，可以在 C·H·米勒和 O·普雷斯勒（O. Pressler）合著的 *Leitfaden der Projectionslehre* 一书[①]中找到。例如，书中讲解了如何准确地画出相切的圆弧、浮标、形形

① 这是一本立体几何作图的练习册，莱比锡，1903 年。

色色的水晶玻璃制品等，并列举了完全不同的生物学领域中的例子，如细胞组织、蜂窝等等。

现在叙述把轴测映射与我们讨论过的仿射变换连接起来的定理：通过平行映射和相似变换而实现从空间到平面的映射（轴测映射），从分析上来说，是利用一个系数行列式为零的仿射变换

$$\begin{cases} x' = a_1x + b_1y + c_1z, \\ y' = a_2x + b_2y + c_2z, \\ z' = a_3x + b_3y + c_3z, \end{cases} \quad \text{其中}\ \Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0。 \tag{7}$$

这正是前面推迟考虑的例外情况。这样，你们就看到了这些"退化变换"的重要性，而遗憾的是这一点常常受到不应有的忽视。其逆也是对的，即每个 $\Delta = 0$ 的这种变换，给出一个轴测映射。确实，这要预先假定，不是变换的所有系数或二阶子行列式都为零，否则就意味着进一步退化。这一点，我现在不予讨论，因为放在以后研究起来比较方便。

为了证明我们的论断，我们暂时假设：由(7)式给出的所有点 x', y', z'（对任意 x, y, z）都真的在一个平面上，即存在 3 个数 k_1, k_2, k_3。使得

$$k_1x' + k_2y' + k_3z' = 0 \tag{8}$$

对 x, y, z 恒成立。据(7)式，此恒等式等价于 3 个齐次线性方程

$$\begin{cases} k_1a_1 + k_2a_2 + k_3a_3 = 0, \\ k_1b_1 + k_2b_2 + k_3b_3 = 0, \\ k_1c_1 + k_2c_2 + k_3c_3 = 0。 \end{cases} \tag{8'}$$

只要行列式 $\Delta = 0$，而 9 个余子式不全为零，这些方程将完全唯一地确定出比 $k_1 : k_2 : k_3$。因此，所有像点 x', y', z' 实际上位于由方程(8′)所确定的平面(8)上。

现在在空间 R' 内作一个新直角坐标系，使得平面(8)成为 $x'y'$

平面($z'=0$)。这样,R 中的每一点必对应于 $z'=0$ 中的一点。于是,在新坐标系中,仿射变换方程必然具有形式

$$\begin{cases} x' = A_1 x + B_1 y + C_1 z, \\ y' = A_2 x + B_2 y + C_2 z, \\ z' = 0。\end{cases} \tag{9}$$

由于最后一行的特殊形式,使得变换行列式始终为零,所以这 6 个系数 A_1, …, C_2 是完全任意的,但应使 3 个子行列式不全为零,即

$$A_1 : B_1 : C_1 \neq A_2 : B_2 : C_2。$$

否则,将出现我们前面排除在外的退化情况。

我现在证明:由上述解析定义的从空间 R 到 $x'y'$平面 E'的映射,与前面定义的轴测映射完全一致。像前面对行列式不为零的仿射变换的讨论那样,我们将阐明变换(9)的主要性质,分几个步骤来加以证明。

(1) 首先,R 中的每一个点 x, y, z,显然对应于 E'中唯一点 (x', y')。反之,若在 E'内给出一点(x', y'),则方程(9)表明,在 R 内的对应的点在两个给定的平面上,而按我们的假设,它们的系数不成比例,因此它们相交为一条直线。在我们的变换下,这条线在所有点必然对应于同样的点(x', y')。如果使点(x', y')变化,则两个平面都将平行于自己而移动,因为系数 A_1, B_1, C_1 和 A_2, B_2, C_2 保持不变。因此它们的交线相互平行,从而得到结论:E'的每一点,对应于 R 内双重无穷个平行直线集之中一条直线上的所有点。这立即指明我们的映射和空间平行投影之间的联系。

(2) 像在讨论一般仿射变换时所说的第(3)点一样,E'内线段的分量对应于R 内的自由矢量 X, Y, Z,现在得出公式为

$$\begin{cases} X' = A_1 X + B_1 Y + C_1 Z, \\ Y' = A_2 X + B_2 Y + C_2 Z, \\ Z' = 0。\end{cases} \tag{10}$$

这再次表明，在 R 内的每个自由矢量，对应于平面 E' 内的自由矢量 X'，Y'，或较严格地说，如果使空间 R 中的一个线段在保持其长度与方向下作平行移动，则在平面 E' 内的对应线段也保持其长度与方向作平行移动。

(3) 现在考虑一个特殊的情形：在 x 轴上从(0, 0, 0)到(1, 0, 0)的单位矢量 $X=1$，$Y=Z=0$。根据(10)式，它对应于 E' 中从原点 O' 到坐标为 (A_1, A_2) 的点的矢量 $X'=A_1$，$Y'=A_2$。完全类似地，y 与 z 轴上的单位矢量，分别对应于从原点 O' 到点 (B_1, B_2) 和 (C_1, C_2) 的两个矢量。在 E' 中的 3 个矢量，我们将简称为 (A)，(B)，(C)（图 15.9）。它们可以任意选择，因为它们的终点坐标决定了仿射变换(9)的 6 个任意参数，所以它们完全确定了映射。现在，这 3 个矢量必须不在同一条直线上，而为了简单起见，我们将假设没有两个矢量处在同一条直线上。其结果如下：R 内各坐标轴上的 3 个单位矢量，被映射成在 E' 内经过原点 O' 的 3 个任意矢量，知道了 3 个任意矢量之后，仿射变换就完全确定了。

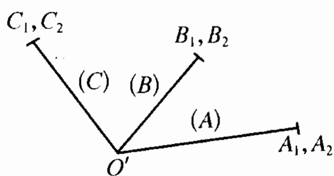


图 15.9

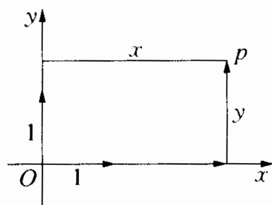


图 15.10

(4) 为了用几何方法求出(A)，(B)，(C)的像，我们从 xy 平面上任意点 $P(x, y, z=0)$ 出发。使 x 轴上单位矢量乘以标量 x，使 y 轴上单位矢量乘以标量 y，然后把乘积加起来，我们得到从 O 到 P 的矢量（图 15.10）。但我们可以立即把这个过程转移到 E' 上，因为 xy 平面与 E' 之间的关系显然是一个通常的二维仿射变换（行列式不为零）。于是，分别对矢量(A)，(B)乘以标量 x，y，并按平行四边

形法则将这两个乘积加起来，我们就得到 P 的图像 P'（图 15.11）。我们可以用这个方法在 E' 中作出任一点的像，从而可逐点地作出 xy 平面内任何图的像。

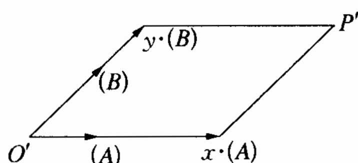


图 15.11

(5) 如果将这些考虑转移到空间 R 中的任意一点（图 15.12）上，可以很容易地证明下列结果：如果将平行四边形法则应用于分量 (A)，(B)，(C) 分别与标量 x，y，z 相乘后之和，我们就得到坐标为 (x, y, z) 的点 P 的图像 P'。因为加法是可交换的，所以可以用 $(1\times2\times3=)6$ 个不同方法来作出这个图，所得的 P'，就是平行与相等线段用 6 个不同方式作加法所得矢量的端点。这样作出的图（图 15.12）显然表示 R 中以 3 个坐标平面及通过点 P 平行于它们的平面为界面的长方体。我们从年轻时起就习惯于把这个平面图当作立体图，特别是用较重的笔画画出前面的各边使其外观发生夸大。这种习惯非常顽固，使得平行六面体的这个图像几乎显得微不足道，而实际上它代表着一个很有价值的定理。

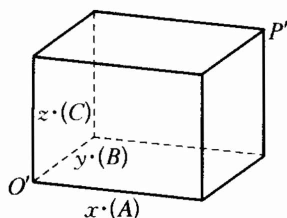


图 15.12

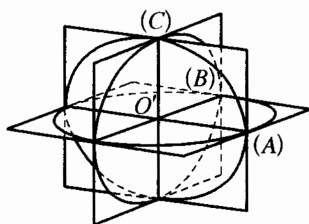


图 15.13

(6) 借助于上述这种作图法，可以在 E' 内作出任何空间图形，即其所有点的像。我只考虑一个例子：如果有一个中心在原点、半径为 1 的球，于是将主要考虑由各坐标平面与它相交的各个圆。在 xy 平面上的截圆，以 x，y 轴上的单位矢量为共轭轴，即为彼此垂直的半径。因为是仿射关系，它将对应于以 O' 为中心、以矢量 (A) 和 (B)

作为共轭半径的一个椭圆，见图 15.13，而椭圆内切于由矢量 2(A) 与2(B)组成的平行四边形。类似地，对应于其他两个截圆的椭圆将以 O'为中心，以(B)，(C)和(A)，(C)为共轭半径。

(7) 我们既已作出一个完整的图像，以表示行列式为 O 的仿射变换(9)的性质，现在就必须采取最后的关键步骤，以说明这些仿射变换确实和我们所断定的那样产生于轴测映射。这主要要求用到所谓波尔克(K. Pohlke)基本定理。K·波尔克是柏林建筑学院的教授，这个定理是他在 1853 年发现并在 1860 年发表于 *Lehrbuch der darstellenden Geometrie* 一书①。H·A·许瓦兹(H. A. Schwartz)并于 1863 年发表了该定理的第一个证明②，同时讲述了有关这个发现的一个有趣的故事梗概，你们应该读一读。

波尔克本人并未用解析法定义轴测映射，而是用几何方法，通过平行射线(加上必要的相似变换)作为一个空间图形而给出定义。他的定理说：在空间 3 个坐标轴上的单位矢量，在这种变换下可变成 E'内通过O'的 3 个任意矢量。而从第(3)点说明中已看出，我们用解析法定义的映射，实际上已导出了 3 个这样的矢量。因此，对我们来说，波尔克定理的基本意义在于：我们用解析法定义的变换(9)是由平行投影和长度单位改变而实现的，而借助于波尔克定理，使第(1)点中所说的平行线变成了投影射线。

(8) 我想大致为这样建立的定理给出一个直接的解析证明。如果我们把注意力集中在 R 内的两个平行平面束

$$A_1x + B_1y + C_1z = \xi,$$
$$A_2x + B_2y + C_2z = \eta,$$

其中 ξ，η 为参变量，于是 ξ，η 的每对值决定了有关平行线中的一条

① 该书分两部分，1876 年柏林第四版。这个定理见第一部分第 109 页。

② H·A·许瓦兹的证明发表于 *Journal für die reine und angewandte Mathematik*，第 63 卷，第 309 页，或见其《数学著作集》，第 2 卷，第 1 页，柏林，1890 年。

平行线。如果有可能将含有适当单位长度的直角坐标系 x'，y'的平面E'放在空间R之内，使得每条射线ξ，η在点$x'=\xi$，$y'=\eta$处贯穿平面E'，则将按希望用几何方法作出了映射。

为此，必须使平面$\xi=0$，$\eta=0$分别在坐标轴$O'y'$，$O'x'$，即相互垂直的两直线处与平面E'相截。如果θ_1，θ_2(确定平面E'的位置)是这些轴与直线$\xi=0$，$\eta=0$之间的角(图 15.14)，且令α为平面$\xi=0$，$\eta=0$之间的角(已知)，则对由$\xi=0$，$\eta=0$与E'形成的三面角应用球面三角形的余弦定理，我们得到$O'x'$，$O'y'$的角的余弦为

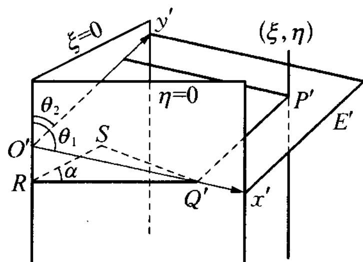


图 15.14

$$\cos\theta_1\cos\theta_2+\sin\theta_1\sin\theta_2\cos\alpha,$$

且当

$$\cot\theta_1\cdot\cot\theta_2=-\cos\alpha \tag{11}$$

时，这个角为直角。

现在，每个平面$A_1x+B_1y+C_1z=\xi$与E'相交于直线$x'=$常量。如果Q'是它与x'轴的交点，则对应的x'值在E'内坐标系的一个待定尺度因子λ的范围内等于$O'Q'$。如果分别作垂直于平面$\xi=0$和直线$\xi=\eta=0$的直线$Q'S$与$Q'R$，则有

$$O'Q'=\frac{Q'R}{\sin\theta_1},\quad Q'R=\frac{Q'S}{\sin\alpha}。$$

又因$Q'S$作为平面

$$A_1x+B_1y+C_1z=0$$

与平面

$$A_1x+B_1y+C_1z=\xi$$

之间的公垂线，是不难借助空间解析几何的已知公式表达的，于是最后得

$$x' = \lambda \cdot Q'O' = \lambda \frac{\xi}{\sqrt{A_1^2 + B_1^2 + C_1^2} \cdot \sin\theta_1 \cdot \sin\alpha}。$$

类似地，可求得 $A_2x + B_2y + C_2z = \eta$ 与 E' 的交点的坐标 y' 为

$$y' = \lambda \cdot \frac{\eta}{\sqrt{A_2^2 + B_2^2 + C_2^2} \cdot \sin\theta_2 \cdot \sin\alpha}。$$

因为我们希望每条由参数 ξ，η 确定的平行射线在点 $x' = \xi$，$y' = \eta$ 处穿过平面 E'，所以必须有

$$\begin{aligned} \lambda &= \sqrt{A_1^2 + B_1^2 + C_1^2} \cdot \sin\theta_1 \cdot \sin\alpha, \\ &= \sqrt{A_2^2 + B_2^2 + C_2^2} \cdot \sin\theta_2 \cdot \sin\alpha。\end{aligned} \tag{12}$$

因此可得 θ_1，θ_2 的第二个方程：

$$\sin\theta_1 \cdot \sqrt{A_1^2 + B_1^2 + C_1^2} = \sin\theta_2 \sqrt{A_2^2 + B_2^2 + C_2^2}。 \tag{13}$$

简单的计算表明，除去符号外，方程组(11)和方程组(13)对 $\cot\theta_1$ 和 $\cot\theta_2$ 只有一对实数解，即本质上只存在平面 E' 的一个位置(当然对于 $\xi = 0$，$\eta = 0$ 的公共法平面)，就我们按方程(12)在 E 的直角坐标系上所选的尺度而言，仿射变换 $x' = \xi$，$y' = \eta$ 在此平面位置上是通过轴测映射实现的。如果从 x'，y' 轴的单位点落在线 $\xi = 1$，$\eta = 0$ 和 $\xi = 0$，$\eta = 1$ 这个条件出发，我们可以用比较几何化的形式提出所述全部论据。这样，问题就在于找到一个与给出的三棱柱截成一个等腰直角三角形的平面 E'。

在这个详细的说明之后，我已无需讨论前面谈过的逆定理，即每个轴测映射表示一个行列式为零的仿射变换。像前面一样，首先利用 R 中 x 轴、y 轴的平行投影所产生的平面 E' 内的斜坐标系，然后通过线性变换过渡到 E' 内原来给出的直角坐标系，即可证明这个逆

定理。

在关于仿射变换的这一章结束时，我想提醒你们注意，可以做一个实验，利用投影灯（必须设想它处于无穷远处）把一些简单的模型（正方形、圆、椭圆、立方体）的阴影图像投射到屏幕上，即得轴测映射图的实例。用这个方法，可以使我们的结果和图形得到证实。特别是，通过模型和屏幕的运动，使 3 个相互垂直棒的阴影图发生各种类型的变化，可以很容易地用实验证明波尔克定理。

现在进入新的一章，它将讨论把仿射变换作为特殊情况包括在内的更一般的变换，即射影变换。

第十六章

投 影 变 换

本章一开始就处理三维空间。

(1) 我从投影变换的解析定义开始。我们不再取 x', y', z' 为 x, y, z 的整函数,而是取为 x, y, z 的有理线性函数,但有一个根本条件,即其分母相同

$$\begin{cases} x' = \dfrac{a_1 x + b_1 y + c_1 z + d_1}{a_4 x + b_4 y + c_4 z + d_4}, \\ y' = \dfrac{a_2 x + b_2 y + c_2 z + d_2}{a_4 x + b_4 y + c_4 z + d_4}, \\ z' = \dfrac{a_3 x + b_3 y + c_3 z + d_3}{a_4 x + b_4 y + c_4 z + d_4}。 \end{cases} \tag{1}$$

只要公分母不为零,则每点 x, y, z 对应于一个确定的有限点 x', y', z'。但是如果点 x, y, z 趋向于平面 $a_4 x + b_4 y + c_4 z + d_4 = 0$,和仿射变换不同的是:其对应点 x', y', z' 将移至无穷远。因此,我们称这种平面为投影平面,它的点为投影点,并说它在投影变换下对应于空间中无穷部分或对应于无穷远点。

(2) 处理这里出现的问题时,正如你们所知道的,用齐次坐标是很方便的,即用按方程

$$x = \frac{\xi}{\tau},\ y = \frac{\eta}{\tau},\ z = \frac{\zeta}{\tau}$$

确定的 4 个量 ξ, η, ζ, τ 来代替点的 3 个坐标 x, y, z。这 4 个量彼此独立变化,但不同时为零,没有一个会变为无穷。因而,每个点 x,

y, z 对应于无穷多组值$\rho\xi$, $\rho\eta$, $\rho\zeta$, $\rho\tau$，其中 $\rho\neq 0$ 是一个任意因子。反之，每组 $\tau\neq 0$ 时，每组值 ξ, η, ζ, τ 都确定一个有限点 x, y, z(所有组 $\rho\xi$, $\rho\eta$, $\rho\zeta$, $\rho\tau$ 都给出同样的点)。当 $\tau=0$ 时，商 x, y, z 中至少有一个变为无穷。于是我们规定，每组值 ξ, η, ζ, $\tau=0$ 表示一个"无穷远点"，且所有组 $\rho\xi$, $\rho\eta$, $\rho\zeta$, 0 表示同一个无穷远点。我们用这种确切的方法引进了"无穷远点"并添加到原有的有限点集内。

经验表明，至少对初学者来说，齐次坐标的运算总是头痛的事。我相信，其原因是任意因子 ρ 带来的这些量的某种不确定与流动性质，我希望这样的说明能减轻这种不适应。

为了同样的目的，作出某些附带的说明，可能有助于理解与齐次坐标相联系的某些几何表示。我首先只谈平面 E 的情况。在这种情况下，我们把直角坐标写成

$$x=\frac{\xi}{\tau},\quad y=\frac{\eta}{\tau}。$$

我们现在把 ξ, η, τ 解释为空间的直角坐标，并在这个空间内，选择平行于 $\xi\eta$ 平面的平面 $\tau=1$ 作为平面E(图 16.1)。在这个平面E上，令$x=\xi$, $y=\eta$。如果用直线把E上的点 x, y 与点 O 连接起来，那么对于这条直线上的点$\frac{\xi}{\tau}$和$\frac{\eta}{\tau}$就是常量，而当 $\tau=1$ 时有 $\xi=x$, $\eta=y$, 故有

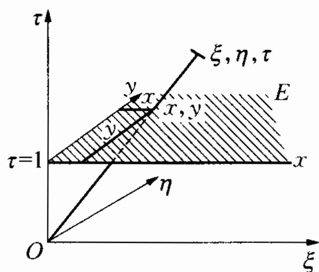


图 16.1

$$\frac{\xi}{\tau}=x,\quad \frac{\eta}{\tau}=y。$$

于是，引入齐次坐标，简单地意味着，从平面 E 到以 O 为中心的空间射线束的映像，而 E 是这射线束的一个截面。一个点的齐次坐标，就成了该点投影射线上所有点的空间坐标。因为 E 的每点对应于射线上的无穷多个点，齐次坐标的不确定的意义就变得清楚了。把

$\xi=\eta=\tau=0$ 排除在外的几何意义在于点 O 不确定射线，因而不对应 E 中的点。又因把 O 与有限点连接，即得到所有射线，所以不需要取 ξ, η, τ 的无穷值。最后，用 $\tau=0$ 给出的、过 O 点的平行射线代替平面 E 的无穷远区域，就使坐标避免了无穷大值。

而且，"无穷远直线"这个常用的术语，也在这里找到了清楚的几何含义。解析地说，就像每条有限直线有一个线性方程一样，它只不过是所有"无穷远点"满足线性方程 $\tau=0$ 的抽象说法而已。但现在，我们可以用几何的语言说，E 内每条直线属于空间中通过点 O 的一个平面上的线束，反之，除去平面 $\tau=0$ 外，空间中通过 O 的每个平面上的线束，都确定 E 内的一条直线。因此，在 E 内指定对应于 $\tau=0$ 的点集为一条直线，似乎是恰当的，因而我们得到"无穷远直线"。

如果对三维空间引入齐次坐标，我们可以构成类似的表示方法。我们把空间设想成四维辅助空间 ξ, η, ζ, τ 的一个截面 $\tau=1$，并把它和从辅助空间的原点投射过来的它的空间线束联系起来。于是，几乎可以逐句引用前面的话，毫无困难地实现所有其他的想法，特别是对无穷远元素作出解释。在这方面，四维空间只是一个方便的表达方法，并未给予什么神秘的意义。

(3) 如果我们将两空间 R 与 R' 的齐次坐标引入射影变换方程(1)，由于它们有公共分母，通过引入一个任意比例因子 ρ，可将其分解成下列 4 个方程：

$$\begin{cases}\rho'\xi'=a_1\xi+b_1\eta+c_1\zeta+d_1\tau,\\ \rho'\eta'=a_2\xi+b_2\eta+c_2\zeta+d_2\tau,\\ \rho'\zeta'=a_3\xi+b_3\eta+c_3\zeta+d_3\tau,\\ \rho'\tau'=a_4\xi+b_4\eta+c_4\zeta+d_4\tau.\end{cases}\tag{2}$$

把任意因子 ρ' 丢开不考虑，我们看到这是最一般的 4 个变量的齐次线性变换。因此，它表示两个四维辅助空间 P, P' 之间的仿射关系，

我们可以用前述第(2)点中的方式来解释其中的齐次坐标。如果仍以平面为限,则所有这些都能给出较具体的表示:我们可以用空间的任意仿射变换来得到平面上最一般的射影变换,只要使通过固定中心 O 的空间投影线束与平面的交点成为原变换的对应点即可。如果加上一个过点 O 的相似变换,我们总可得到对应于因子 ρ' 的同一投影。因为这个变换把每条过点 O 的射线变为自己,而投影性质只依赖于这些射线与平面的交。

前面利用辅助空间 P, P' 时所遵循的原理,称为投影与截面原理。这个原理非常有用,因为一般来说,它利用 $n+1$ 维辅助空间,使 n 维空间中的复杂关系简化并易于理解了。

(4) 现在我们对 ξ, η, ζ, τ 解变换方程组(2)。行列式理论证明,ξ, η, ζ, τ 是 ξ', η', ζ', τ' 的类似的线性齐次组合,当然也带有比例因子 ρ:

$$\begin{cases}\rho\xi = a_1'\xi' + b_1'\eta' + c_1'\zeta' + d_1'\tau',\\ \rho\eta = a_2'\xi' + b_2'\eta' + c_2'\zeta' + d_2'\tau',\\ \rho\zeta = a_3'\xi' + b_3'\eta' + c_3'\zeta' + d_3'\tau',\\ \rho\tau = a_4'\xi' + b_4'\eta' + c_4'\zeta' + d_4'\tau'。\end{cases} \tag{3}$$

只要行列式

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix}$$

不等于零。因此,值组 ξ, η, ζ, τ 与 ξ', η', ζ', τ' 之间是一一对应的(在相应任意公共因子范围内)。

然而,正如你们预期那样,从我们对仿射变换中 $\Delta = 0$ 的情形的处理可以看出,$\Delta = 0$ 在这里也是有特别意义的,不能被忽略。像每个中心投影(例如照相)一样,它代表空间到平面的映射。但现在,我

们将考虑 $\Delta \neq 0$ 的一般情形。

(5) 从方程组(2)和方程组(3)立即推出，当 ξ, η, ζ, τ 之间存在线性关系时，ξ', η', ζ', τ' 之间也存在线性关系，反之亦然。所以，每一个平面对应一个平面，特别是 R' 内的无穷远平面，一般对应于 R 内的一个有限平面，即前面所说的化零平面。因此，无穷远平面的引出已达到了方便的目的，因为它可以使前述定理毫无例外地成立。进而推出，每条直线对应一条直线，用莫比乌斯的话来说，每个投影变换是一个共线变换。

(6) 现在应该指出，上述定理之逆也是对的：空间上每个共线变换，即使直线对应于直线，且满足其他某些几乎不言而喻条件的可逆单值变换，是一种投影变换，即由方程(1)或方程(2)确定。

为了方便起见，这里只对平面上的投影变换给出莫比乌斯的证明。对空间上的投影变换，可以类似地证明。对平面的证明如下：从一个任意的共线变换中，选出对应的两个 4 点组，我们将证明：(a) 始终存在着把这两个 4 点组相互变换的投影变换；而一个投影变换也是共线的；(b) 在某些条件下，使两个 4 点组相互对应的只有一个共线变换。因此，这个投影变换必然与给出的共线变换相同，从而证明了这个定理。现在对这两步证明作出详细的说明。

(a) 我们注意到，在平面内的投影变换方程

$$\rho'\xi' = a_1\xi + b_1\eta + d_1\tau,$$
$$\rho'\eta' = a_2\xi + b_2\eta + d_2\tau,$$
$$\rho'\tau' = a_3\xi + b_3\eta + d_3\tau$$

含有 9－1＝8 个常数(ρ'的改变不影响变换)。两个给定的点在一个射影变换中相互对应，只需要投影变换的常数满足两个线性方程，因为我们只涉及 3 个齐次坐标的比。因此，两个 4 点组的对应，表示 2×4＝8 个线性条件，或更准确地说，表示 9 个量 a_1, …, d_3 之间的 8 个线性齐次方程。你们知道，这种方程组总是有解的。因此，用这种办法找出了把给出的两组 4 个点作相互变换的投影变换的所有常

数。事实上，只要所给出的 4 点组处于“一般位置”，即 4 点中任意 3 点都不共线，就能保证这是一个行列式不化为零的“严格的”投影变换。这是我们的定理所唯一要求的情形。

(b) 现在设平面 E 和 E' 之间有一个任意的共线变换，1，2，3，4 是平面 E 上的任 3 点都不共线的 4 点，$1'$，$2'$，$3'$，$4'$ 是满足同样条件的 E' 上的对应点。我们的论断是这个共线变换完全由这两个 4 点组之间的对应所确定。我们将给出这个证明，因为这个共线变换可以用一个，且仅能用一个方法由这两个对应的 4 点组建立起来，即仅仅借助于 4 点组的两个特性(唯一性和直线间的相互对应)。我们的主要工具就是与蜘蛛网类似的所谓莫比乌斯网。开始，我们在每个平面上各作 6 条直线，连接 4 个点(图 16.2)。它们在这个共线变换下必然相互对应，因为直线 1 2 必然对应于 E' 内的一条直线，此直线必然包含作为 1 的像的 $1'$ 和作为 2 的像的 $2'$，而它只能是直线 $1'2'$。类似地，对应直线的各交点必然相互对应，例如点(1，4，2，3)与点($1'$，$4'$，$2'$，$3'$)对应：这从共线性和唯一性即可推出。如果用直线与新的点连接，把这些直线延长到与原来的直线相交，再把这些交点连接起来，循此以往，在每个平面上将出现越来越稠密的线与点组成的网。这些点和线在所要求的共线变换下，必然成对地相互对应。

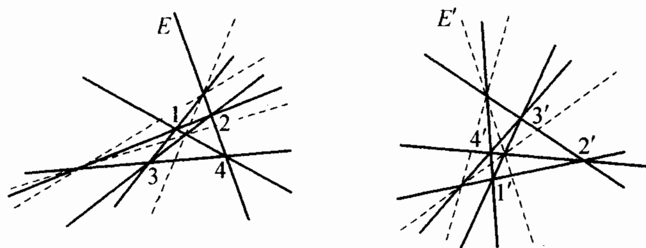


图 16.2

如果现在在 E 内任选一点，例如说它是网的一个角点，或者不是，则可以把网作得任意小，而把这一点包围在一个网眼内，即使其

成为角点的极限点。在前一种情形下，E'中的点作为对应的网的角点而唯一被确定。至于第二种情形，则必须对其线性的定义加上莫比乌斯认为是不言而喻的一个条件，即此映射必须是连续的，即E内一个点集的每个极限点对应于E'中对应点集的极限点。据此及前面的说明，也可以推出，在第二种情况下，E'中的对应点也是唯一确定的。于是，在共线变换是连续的情况下证明了论断(6)的正确性。用同样的方法，可以证明，通常空间里连续的共线变换，由5个对应的点对确定；在n维空间中，则由$n+2$个对应的点对确定。

现在回到第(6)点开头的讨论。结果得出下面这个准确的定理：投影变换是实现直线到直线的唯一连续可逆变换。

绕了这些弯后，请让我们回到第(5)点中提出的研究，即在投影变换下或共线变换下各基本几何流形的性质。我们在那里已经讲过，一个无限平面或直线可以被映射成一个同类的图形，所以这些概念对于投影变换有着确定不变的意义。在这个性质上，一般的投影变换与仿射变换相同，但在平行性质方面不相同。

(7) 平行性概念所造成的情况。确实，两条直线的平行在投影变换下不一定像仿射变换下那样保持着。相反，空间的无穷远平面可以对应于任何有限平面(化零平面)，而两条平行线的公共无穷远点则一般将对应于化零平面上的一个有限点，而对应于两条平行线的两条直线则在此平面上相交。借助齐次坐标，可严格证明这一点。事实上，在这里也可以看到，平行性概念并没有被完全打乱，而是变成了一个完全确定的一般概念的一部分。空间的无穷远点组成一个平面，它可通过投影变换变成空间任何其他(有限)平面，在这个意义上它与所有平面具有同等地位。用“无穷远”这个词来描述只是在某种程度上刻画出它的任意性。于是，称各直线(或平面)是相互平行的，如果它们在这个特殊(无穷远)平面相交的话。通过投影变换，它们可以被变成在其他固定平面相交的直线(或平面)，此时新的直线(或平面)就不再平行了。

与这个性质相联系的事实是，在投影变换下，格拉斯曼的基本流形将相应地失去了不变量的意义。自由矢量绝不会变为另外的自由矢量，滑动矢量亦然，等等。让我们看一看在空间 R 具有 6 个坐标：

$$X = x_1 - x_2,\ Y = y_1 - y_2,\ Z = z_1 - z_2,\ L = y_1 z_2 - y_2 z_1,$$

$$M = x_2 z_1 + z_2 x_1,\ N = x_1 y_2 - y_1 x_2$$

的一个线段，并设在射影变换(1)

$$x_1' = \frac{a_1 x_1 + b_1 y_1 + c_1 z_1 + d_1}{a_4 x_1 + b_4 y_1 + c_4 z_1 + d_4} \text{ 等等,}$$

$$x_2' = \frac{a_1 x_2 + b_1 y_2 + c_1 z_2 + d_1}{a_4 x_2 + b_4 y_2 + c_4 z_2 + d_4} \text{ 等等}$$

下，由对应于点$(x_1,\ y_1)$和$(x_2,\ y_2)$的点$(x_1',\ y_1')$和$(x_2',\ y_2')$构成的类似量为 $X', \cdots, N'$。通过这些公式，$X', \cdots, N'$变成分式，其分子只是 6 个量 $X, \cdots, N$ 与常数系数的线性组合，而对它们相同的分母 $(a_4 x_1 + b_4 y_1 + c_4 z_1 + d_4)(a_4 x_2 + b_4 y_2 + c_4 z_2 + d_4)$，包含不能只用 $X, \cdots, N$ 表达的点坐标本身。因此，变换后的线段的坐标，不仅依赖于原来线段的坐标，而且依赖于端点的位置。如果沿线段所在直线使之滑动，则 $X, \cdots, N$ 不会改变，而 $X', \cdots, N'$一般将改变，即$(1,\ 2')$不是格拉斯曼意义上的一个线段。

然而，无限直线在投影变换下仍保持这个性质，因为它是由比例 $X' : Y' : \cdots : N'$确定，出毛病的公分母被约去。因此，这个比实际上只由比 $X : Y : \cdots : N$ 表达。

(8) 仍然有某些重要的流形在投影变换下保持为同类流形。首先，通过乘以公分母 $a_4 x + b_4 y + c_4 z + d_4$ 的平方，可以看到，每一个 $x,\ y,\ z$ 的二次方程变换成 $x',\ y',\ z'$的二次方程，反之亦然。这表明，在空间 R 中的每一个二次曲面对应于 R'中的二次曲面。因此，每个这种曲面与平面的交线，即在空间 R 内的每个二次曲线对应于 R'中的二次曲面。同样，用一个或几个坐标的方程所确定的任何代

数流形，将变换为同类型的代数流形。因此，这些流形的性质是射影变换下的不变量。

(9) 除了这些用方程确定的不变流形外，我必须指出在所有投影变换下保持不变的一个数值量，它取代了距离和角度的概念。正如你们所知道的，距离和角度即使在仿射变换下它也不是不变量，更不必说投影变换了。首先谈直线，我们考虑任选 4 点 1，2，3，4 间的距离的某个函数，即以前说过的交比

$$\frac{\overline{12}}{\overline{14}}:\frac{\overline{32}}{\overline{34}}=\frac{\overline{12}}{\overline{14}}\cdot\frac{\overline{34}}{\overline{32}},$$

不难通过计算验明这个量在投影变换下的不变性。后面讨论不变量理论时，我们将这样做。

如果用角的正弦来取代角，则情况对于射线束也十分类似。如果 1，2，3，4 是束中的射线或平面，则其交比表达为

$$\frac{\sin(1,\ 2)}{\sin(1,\ 4)}:\frac{\sin(3,\ 2)}{\sin(3,\ 4)}=\frac{\sin(1,\ 2)\sin(3,\ 4)}{\sin(1,\ 4)\sin(3,\ 2)}。$$

因为这些交比是投影变换下首先发现的数值不变量，所以许多研究投影几何学的人认为把所有其他的不变量化为交比，即令这种转化是十分人为的，也是值得称道的目标。稍后，我们将比较彻底地考虑这些问题。

这几点说明已足以表明，如何根据投影变换下的性质而严格地区分不同的几何概念。在这种变换下保持不变的一切都是投影几何的研究对象。我已经讲过，投影几何是 19 世纪产生的一门学科。这个学科名称现在已十分通用，它比早先常用的“位置几何”这个名称好。以前数学家用位置几何这个名称，是想拿它来同度量几何或初等几何对比，而把一切几何性质，包括在投影变换下改变的性质都列入度量几何或初等几何的范围。旧的名称完全没有考虑到许多度量性质，特别是交比值。

我现在要像讨论仿射变换时一样，讨论射影变换的应用。

(1) 我只从画法几何出发，讨论几个典型的例子而不企求系统化。

(a) 第一个例子是通过中心投影把空间映射到平面的例子。这是轴测(平行)映射的直接推广。这里，投影射线是一个有限点而不是从无穷远点发出的。我们选坐标原点为投影中心，以 $z=1$ 为投影平面(图 16.3)。于是，任意点 $P(x, y, z)$ 的像 $P'(x', y', z')$ 均有 $z'=1$，且因 P 与 P' 处在同一条通过 O 的射线上，故有

$$x':y':z'=x:y:z。$$

因此，我们的映射方程是

$$x'=\frac{x}{z},\ y'=\frac{y}{z},\ z'=\frac{z}{x},$$

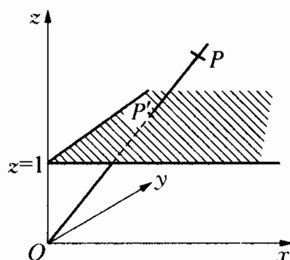


图 16.3

这是一个特殊的投影变换。由于它与轴测映射相似，自然使我们怀疑它的行列式等于零。事实上，过渡到齐次坐标，我们得

$$\rho'\xi'=\xi,\ \rho'\eta'=\eta,\ \rho'\zeta'=\zeta,\ \rho'\tau'=\tau,$$

代换的行列式为

$$\Delta=\begin{vmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&1&0\end{vmatrix}=0。$$

只要你们注意到，一般来说，每个平面通过一个行列式不为零的(二维)投影变换。与投影平面相联系，你们就可以用早先讨论中用过的类似方法，很快推出这个变换的各种性质。由此特别推出，一条直线上任意 4 点的交比，或过一点的任何 4 条直线的交比，经过变换是不

变的。

(b) 第二个例子是关于把中心投影作为极限情形包含在内的，具有不为零的行列式的所谓立体投影的投影性质。一个物体的立体像，投入一定远的观察者眼中的射线，与原物体投入相应位置观察者的射线相同。这意味着，在适当选择坐标系方向的情况下，原来点及其像点将在过原点的同一条射线上

$$x' : y' : z' = x : y : z。 \tag{4}$$

这一种情况与前一种情况的差别在于：原来点不是被映射到一个平面上，而是被压缩到某个有限宽度的狭窄空间块内。

我断定，这个变换由公式

$$x' = \frac{(1+k)x}{z+k},\ y' = \frac{(1+k)y}{z+k},\ z' = \frac{(1+k)z}{z+k} \tag{5}$$

来完成。首先，它至少给出了一个投影变换，也显然满足方程(4)。我们利用其对应的齐次方程

$$\rho'\xi' = (1+k)\xi,\quad \rho'\eta' = (1+k)\eta,$$
$$\rho'\zeta' = (1+k)\zeta,\quad \rho'\tau' = \zeta + k\tau$$

求得其行列式为

$$\Delta = \begin{vmatrix} 1+k & 0 & 0 & 0 \\ 0 & 1+k & 0 & 0 \\ 0 & 0 & 1+k & 0 \\ 0 & 0 & 1 & k \end{vmatrix} = k(1+k)^3。$$

除 $k=0$ 或 $k=-1$ 外，它显然不为零。

对 $k=0$ 情形，(5)式正好转变为前面的中心投影公式，即由立体完全退化到平面。$k=-1$ 时给出 $x'=y'=z'=0$，即空间的每个点均由原点代表，这显然是没有用处和无价值的特殊情况。

为了确定起见，我们选择 $k>0$。为了明确变换(5)的几何意义，我们首先注意到，每个平面 $z=$ 常量，变换成一个平行平面

$$z'=\frac{(1+k)z}{z+k}。\tag{6}$$

通过由 O 发出的射线使两个平面相互变换是完全清楚的，现在只需要解释这个规律(6)式。

当 $z=\infty(\tau=0)$ 时，$z'=1+k$。这个与 xy 平面平行且与其距离为 $1+k$ 的平面，是像空间的化零平面，同时在某种意义上形成立体图的背景。物体的空间图在无穷远处的背景被映射而呈现在此平面上。平面 $z=1$ 也起重要作用，因为在这个平面上物与像重合，这是由于当 $z=1$ 时，z' 也等于 1。如果令 z 从 1 增加到 ∞，则 z' 从 1 单调地增加到 $1+k$，即如果将物体限制在平面 $z=1$ 的后面，我们实际上得到一个有限深度 k 的立体像。实际上，能够且始终必须有这样的限制(图 16.4)。

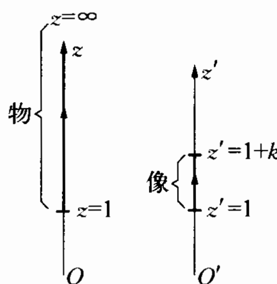


图 16.4

再注意一下(6)式，对点 z，1，z'，0 的交比找到关系式

$$\frac{z-1}{z-0}\cdot\frac{z'-0}{z'-1}=\frac{z-1}{z}\cdot\frac{(1+k)z}{k(z-1)}=\frac{1+k}{k}。$$

一般来说，这两个彼此对应的点 z 与 z' 和点 1 与 0 组成的交比为常数。

我们收集一数学模型，可以表示立方体上的球、旋转圆锥及旋转圆柱的立体投影。从适当的距离处观察，模型对原物体给出了一个十分清楚的形象，当然，心理反应起了重要作用。同样的光线进入人眼这个孤立的事实，并不足以确定立体感，习惯必然起某种重要作用。事实上，因为我们多看立方体上的球，少看狭长六面体上的狭长

的椭球（这正是立体投影图像的情况），所以我们从一开始就不免把光的印象同前一个来源的印象联系起来。对这个反应的进一步考察，可以留给心理学家去进行。

以上种种足以使你们对投影变换在画法几何上的应用获得初步印象。当然，这些定理要求进一步的考虑，在结束这一节以前我不能不鼓励你们彻底研究画法几何，因为我认为这种研究对每一名数学教师是不可缺少的。

（2）我希望讨论的有关投影变换的第二个应用，是如何由此推导几何定理和观点。你们会回忆起，我们曾为类似的目的对仿射变换进行讨论。

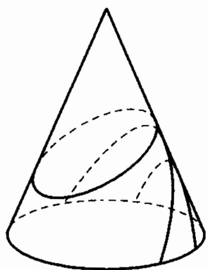


图 16.5

（a）我们从这样的定理出发，即当一个圆通过射影变换或中心投影变换时，它变成一个“圆锥截线”，即圆周上点的射线所组成的圆锥面与一个平面的交线。我有一个模型，可以用来表示如何用这个方法形成一个椭圆、一条双曲线或一条抛物线（图 16.5）。

（b）由此推出，对投影几何而言，只有一种圆锥曲线。因为任何两种圆锥曲线都可变换成一个圆，从而可彼此互相变换。从这个观点来看，区分成椭圆、抛物线和双曲线，并未说明本质差异，只反映了相对于通常所谓“无穷远”直线的不同位置而已。

（c）现在对圆锥曲线推导出下面的基本交比定理：已给一条圆锥曲线上的任意 4 个固定点 1, 2, 3, 4，从同一曲线上第五个可移动点 P 向这 4 点作射线，那么 4 条射线的交比与点 P 的位置无关（图 16.6）。

为了证明这一点，我们回过来看一下通过中心投影产生所述圆锥曲线的那个圆。只要证明在圆的情形下交比不变，一般定理即获证明。设在圆上有 4 个对应点 $1'$, $2'$, $3'$, $4'$，由圆上其他两个任意

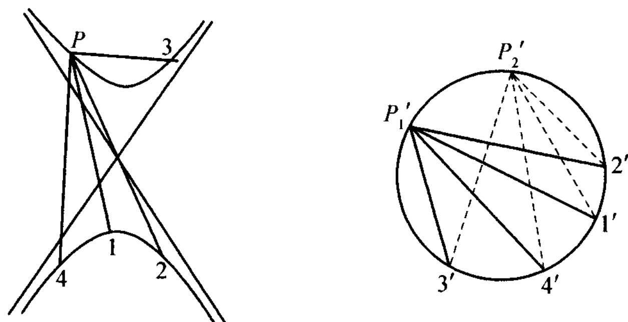


图 16.6

点 P_1'，P_2'向其各作 4 条射线，则两组射线的交比相等。但这是显而易见的，因为按圆周角定理，射线束 $P_1'(1', 2', 3', 4')$与 $P_2'(1', 2', 3', 4')$所构成的角分别对应相等，故这两组射线所成角的正弦的交比当然相等。

(d) 斯坦纳的圆锥曲线实际上是建立在这个定义的基础上，因为他是从两组"射影相关的"射线束出发的，而在这两组射线中，两组对应的 4 条射线有相同的交比。所以，圆锥曲线是两组对应射线的交点的轨迹。

这几点说明已足以使你们明白射影变换对圆锥曲线理论的重要意义。你们可以在任何一本投影几何教材中找到更完善的说明。

现在进入与本章有关的更广泛的内容，讨论不属于至今所考虑过的线性变换范围以内的新的几何变换，从位移转入最一般的几何投影。

第十七章

高阶点变换

现在将研究不是由线性函数，而是由高阶有理函数，甚至是超越函数

$$x' = \phi(x,\ y,\ z),\ y' = \chi(x,\ y,\ z),\ z' = \psi(x,\ y,\ z)$$

所表示的变换。

按本讲义的计划，我不给出系统论述，只介绍一系列在纯数学尤其在数学应用上有一般意义的特别例子。

§17.1 反演变换

本变换将每点 P 变成连接 P 与原点 O 的直线上的点 P'，使 $OP \cdot OP'$ 等于给定常量(图 17.1)。

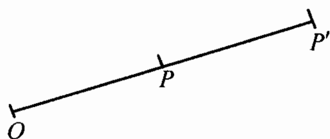


图 17.1

正如你们所知道的，这个变换在纯数学，特别在复变函数理论里起到十分重要的作用，但在物理学和其他的应用中较少出现。后面将充分讨论常数为 1 的特殊应用情况。

(1) 在处理这个变换时，仍从推导它在直角坐标系的方程出发。因点 P 与点 P' 在同一条通过点 O 的直线上，故有

$$x' : y' : z' = x : y : z。 \tag{1}$$

又为了简单起见，设有关常数为 1，从距离 OP 与 OP' 的公式，可得

$$(x^2+y^2+z^2)(x'^2+y'^2+z'^2)=1。\tag{2}$$

因此,变换方程为

$$x'=\frac{x}{x^2+y^2+z^2},\ y'=\frac{y}{x^2+y^2+z^2},\ z'=\frac{z}{x^2+y^2+z^2}。\tag{3}$$

同样可得

$$x=\frac{x'}{x'^2+y'^2+z'^2},\ y=\frac{y'}{x'^2+y'^2+z'^2},$$
$$z=\frac{z'}{x'^2+y'^2+z'^2}。\tag{4}$$

因此,不仅 P' 的坐标由 P 的坐标有理地表达,P 的坐标也由 P' 的坐标有理地表达,而且在两种情况下所出现的函数相同。分母都是一个二次表达式。这里是一个所谓二次双有理变换的特殊情形。而且,存在着正反均为有理函数表示的、更广泛的一类双有理变换(一般是单值可逆)。它们以克里蒙那(Cremona)变换为名,成为一个广泛发展的理论的对象,至少应在讨论其中最简单的一种变换时加以提及。

(2) 方程(3)和方程(4)表明,空间中每一点 P 对应于一点 P',反之,如除去原点(目前),则每点 P' 也对应一点 P。但如令 x, y, z 同时趋向于零,则方程(3)的分母是比分子高阶的无穷小,故 x', y', z' 变成无穷大。因此,可以称原点为变换的零点。反之,如 x', y', z' 以任意方式变成无穷大,按(4)式,x, y, z 都将趋向于零。因而,用早先的术语,可以说一个点对应于整个无穷远平面。但这个“无穷远平面”只是为了适应投影变换而采用的方便说法。它意味着,在这种变换下,空间的无穷远部分在性质上和平面一样,即它将变换成某一有限平面上的点,这样可使定理排除例外,不必引入几种情况。这并不妨碍我们在这里采用不同的表达方式,使这里的定理同样毫无例外地成立。据我们的变换,空间的无穷远处被变换成一个点。

因此,我们简单地说,只有一个无穷远点,它在我们的变换下对应于坐标原点。于是,我们的变换事实上成为无例外的单值可逆变换。

这里也像在我们前面的说明中一样,必须强调:关于无穷远的真实性质,我们所考虑的,一点也没有形而上学的意思。当然总是有人由于习惯上偏向于这一种或那一种的表达形式而希望对无穷远赋予超越的意义。这类两种观点的拥护者有时陷于争论不休的情况。当然,两者都是错误的。他们忘记了我们真正关心的是为了适合某一个目的而作出的任意约定。

(3) 我们的变换的主要性质,一般地说是把球转变为球。事实上,球面方程具有形式

$$A(x'^2+y'^2+z'^2)+Bx'+Cy'+Dz'+E=0。\qquad(5)$$

用(3)式的值 x', y', z'代入,以 $x'^2+y'^2+z'^2$ 除全式,借助于(2)式,可得 $A+Bx+Cy+Dz+E(x^2+y^2+z^2)=0$,这确实是一个球面方程。事实上,应该指出,方程(5)也包括平面(当 $A=0$ 时),在这里可以把它当作一个特殊的球面,而事实上是包含无穷远点的球面。在我们的变换下,它们变成通过与无穷远点对应的点即原点的球。反之,任何经过原点的球变成通过无穷远点的球面,即平面。有了这个规定,球对应球的定理就毫无例外地成立了。

由于两个球面(一个球面和一个平面也同样如此)相交于一个圆,因此也导出圆总是对应于圆,在此,直线是作为"通过无穷远点的圆"而包含在内的。反之,在我们的变换下,一条直线对应于一个通过原点的圆。

(4) 如果以平面的变换为限,上述定理当然仍然成立。例如直线运动问题,这是一个非常初等的问题,而实际上属于非数学家的兴趣范围,但是有了上述定理,才使这个问题得到了一个漂亮的解答。问题是借助一个链杆系统控制一点沿直线运动。从前在制造蒸汽机时,特别重要的是设计一种机械,使活塞的直线运动转换成曲柄的圆周运动。

这使我们对法国军官波塞利(Peaucellier)于1864年制造的“逆转器”产生了兴趣。这个机械装置轰动一时,尽管结构是十分简单而明显的。此装置由6个连杆组成(图17.2)。两条长为l的杆连接在固定点O上;另外4条长为m的杆形成一个菱形,其一对顶点是杆l的端点。设菱形的自由顶点为P与P'。这个装置有两个自由度:首先,可使两个杆l任意相互倾斜;其次,可使这两个杆一起绕点O旋转。但每当发生这种运动,OPP'仍在一条直线上,这是不难用几何方法证明的,且积

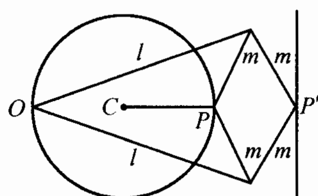


图 17.2

$$OP \cdot OP' = l^2 - m^2 = \text{常量}$$

与P的位置无关。因此,这个装置实际上实现了以点O为中心的一种反演变换。我们只需要使点P沿一个过点O的圆移动,即可迫使点P'(按上述第(3)点中的定理)沿一条直线移动。如果点P附在第七条杆PC上,其另一端点C固定在点O与点P的初始位置的中点,即可产生这种结果,这样就只剩下一个自由度,即点P'将沿一条直线移动。应该指出,点P'不可能遍历整个无限直线,由于给定的杆长不允许过分移动,它的自由移动将受到它到点O的距离不得超过$l+m$的限制。在某些模型中,点C有小小的移动,因而点P所经历的圆只是接近点O,故点P'不是在一条直线上而是在一个半径大的圆周上移动。这个装置的这种用法也可能常常会用到[①]。

(5) 在反演变换的一般性质中,我最后要强调的是保角性。这意味着,两曲面在其交线上任一点所构成之角在变换后仍保持不变。

---

① 还可参考A·B·肯普(A. B. Kempe)的《怎样画直线》,伦敦,1877年,以及 *Naturwissenschaftlich-medezinischen Abhand lungen der Württembergischen Gesellschaft zur Förderung der Wissenshaften*, Abteilung Tübingen,第6分册,1924年。

因为我不涉及这方面的细节，故略去其证明。

(6) 球极平面投影在应用中也起到很重要的作用，可以把它看做是反演变换中特殊的一章。它的获得如下所述：设考虑一个球面，经反演变换变成固定平面 $z'=1$。根据(3)式中第三式，这个球的方程是

$$1=\frac{z}{x^2+y^2+z^2}。$$

可将其写成

$$x^2+y^2+\left(z-\frac{1}{2}\right)^2=\frac{1}{4}。$$

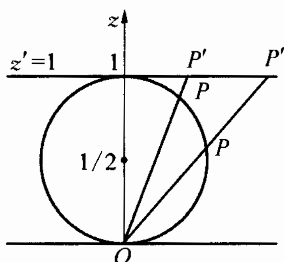


图 17.3

因此，这个被变换成平面 $z'=1$ 的球，半径为 $\frac{1}{2}$，其中心在 z 轴上点 $z=\frac{1}{2}$ 处。它经过原点且与像平面 $z'=1$ 相切(图 17.3)。如果利用通过中心 O 与对应点的空间射线束，那就可以立即把球面与平面的关系弄清楚，并发现相应的点。我们将不加证明地叙述下列定理：

(1) 如果把平面的无穷远部分当作一个点，因而被映射为球面的点 O，则此映射是全平面上单值可逆的。

(2) 球面上的圆对应于平面上的圆，特别是通过点 O 的圆对应于过无穷远点的圆，即直线。

(3) 两个曲面上的对应关系是保角的，或习惯地说，此变换是保角的。

你们当然知道，这个球极平面投影在复变函数论中具有十分重要的意义，我在去年冬天的讲座中经常利用它①。关于其他方面的同等重要的应用，我要提到地理学与天文学。古代天文学家已经知道球极平面投影，即使今天，你们也会在地球的半球和两极区域的地

① 见第一卷第六章第二节。

图上看到球极平面投影。

下面对最后所说的应用领域举若干个例子。

§ 17.2　某些较一般的映射投影

讲到这里，我觉得可以先讲一段闲话。首先，地理绘图理论是中学的一门重要课程。地图是怎么画的，会使每个中学生感兴趣。数学教师如果能介绍一些这方面的知识，就比只讲抽象问题要生动。因此，未来的教师都应了解这个数学领域，而且这个领域能为数学家提供一些有趣的、有关点变换的例子。

如果一开始就把地球想象成从南极到 xy 平面的球极平面投影，那对讲清这方面的问题会有极大的好处。这样的话，相对于那一个极点，任何投到 $\xi\eta$ 平面上的其他映射，将由两个方程 $\xi=\phi(x, y)$，$\eta=\chi(x, y)$给出。

实践中用得最多的是保角映射。如果我们把复变量 $\xi+\mathrm{i}\eta$ 看成是复变量 $x+\mathrm{i}y$ 的一个解析函数

$$\xi+\mathrm{i}\eta=f(x+\mathrm{i}y)=\phi(x, y)+\mathrm{i}\chi(x, y),$$

则如复变函数论中所述，即得到这些映射。但我要强调，地理学中常常用到的，恰恰是不保角映射。因而，保角映射不应像通常那样被看成是唯一重要的映射。

在保角映射中占有显著地位的是数学家默卡多(G. Mercator)于 1550 年左右发现的所谓默卡多投影。每一本地图册里的地球投影，用的都是默卡多投影。

默卡多投影选择对数函数为解析函数。它由方程 $\xi+\mathrm{i}\eta=\log(x+\mathrm{i}y)$ 给出。

作为数学家，我们能立即从这个简短的公式推出投影的性质，但对没有数学训练的地理学家，默卡多投影的处理当然较为困难。在 xy 平面引入极坐标系(图 17.4)，即令 $x+\mathrm{i}y=re^{\mathrm{i}\varphi}$，则有

$$\xi + \mathrm{i}\eta = \log(r \cdot \mathrm{e}^{\mathrm{i}\varphi}) = \log r + \mathrm{i}\phi。$$

故 $\xi = \log r$，$\eta = \phi$。

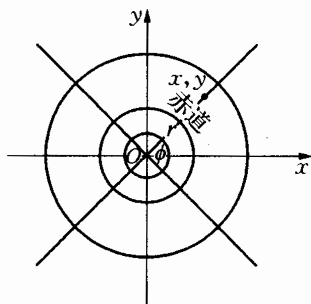


图 17.4

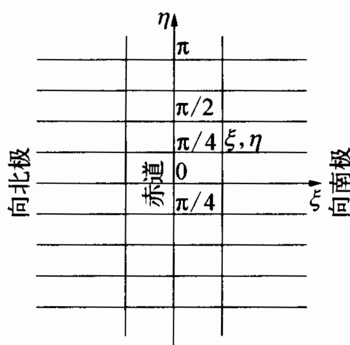


图 17.5

我们取地球的南极为地极平面投影的中心，则 xy 平面的原点 O 对应于地球北极，而射线 $\phi=$ 常量在 xy 平面上对应于经线。于是，在默卡多投影里(图 17.5)，经线变为 $\eta=$ 常量，即平行于 ξ 轴。其上的北极 ($r = 0$，$\xi = -\infty$) 和南极($r = +\infty$，$\xi = +\infty$)，分别位于左侧和右侧无穷远点。因为角 ϕ 未确定，可能相差 2π 的倍数，所以这个映射有无穷多值，每条平行于 ξ 轴、宽为 2π 的带，给出了整个地球表面的映像。$r=$ 常数的纬线圆在默卡多映射中变成平行线 $\xi=$ 常数，也就是说，由于角理所当然保持不变，因此它们是经线映像的正交轨线。赤道 ($r = 1$) 对应于 η 轴($\xi = 0$)。

这个例子可能促使你们进一步研究地理学中的许多映射变换，不过我现在要注意讨论其中一个更一般的定理。你们当中钻研过地理学的人，必然听说过蒂索(Tissot)[①]定理，从我们的观点出发，是很容易把它的内容搞清楚的。

---

① 蒂索。他在 *Die Netzenentwürfe geographischer karten nebst Aufgaben über Abbildungen beliebiger Flächen auf einander* 一书中发表了这个定理，该书由哈默译成德文，1887 年出版于斯图加特。

设在 xy 平面上和 $\xi\eta$ 平面上各有一个表示地球表面的地图，其中每一个的映射都可以是任意的，且不必是保角的。两个映射彼此有某种关系，可将其写成形式 $\xi=\phi(x, y)$，$\eta=\chi(x, y)$。

我们考察两对应点 (x_0, y_0) 和 (ξ_0, η_0) 的邻域，其中 $\xi_0=\phi(x_0, y_0)$，$\eta_0=\chi(x_0, y_0)$。为此，我们用方程

$$
\begin{aligned}
x &= x_0+x', \quad & y &= y_0+y', \\
\xi &= \xi_0+\xi', \quad & \eta &= \eta_0+\eta'
\end{aligned}
$$

导出新变量 (x', y') 和 (ξ', η')。按泰勒定理展开，可得

$$
\xi' = \left(\frac{\partial\phi}{\partial x}\right)_0 \cdot x' + \left(\frac{\partial\phi}{\partial y}\right)_0 \cdot y' + \cdots,
$$

$$
\eta' = \left(\frac{\partial\chi}{\partial x}\right)_0 \cdot x' + \left(\frac{\partial x}{\partial y}\right)_0 \cdot y' + \cdots,
$$

其中导数取于点 $x=x_0$，$y=y_0$，高阶项则由省略号表示之。我们现在以一个充分小的邻域为限，使列出的线性项对实际值 (ξ', η') 给出足够的近似。这当然意味着把没有这种邻域的奇点 (x_0, y_0) 排除在外。因此，所有 4 个偏导数同时为零的点被除外，从而使线性项给出一个有用的近似。于是，如果只考虑这样得到的，在 (x', y') 与 (ξ', η') 之间的线性方程，我们立即得到组成蒂索映射基础的基本定理：两个同一地域的地图，在一个非奇点的邻域内近似地由一个仿射变换联系起来。如果应用前面关于仿射变换的定理，实际上就得到了所谓蒂索定理的全部结果。

我只提醒你们注意几个要点。我们知道，一切都依赖于仿射变换的行列式，即行列式

$$
\Delta = \begin{vmatrix} \left(\dfrac{\partial\varphi}{\partial x}\right)_0 & \left(\dfrac{\partial\varphi}{\partial y}\right)_0 \\ \left(\dfrac{\partial\chi}{\partial x}\right)_0 & \left(\dfrac{\partial\chi}{\partial y}\right)_0 \end{vmatrix}
$$

被称为在点 $x = x_0$, $y = y_0$ 处函数 ϕ 与 χ 的函数行列式。在应用中,我们总是避免 $\Delta = 0$ 的情形,因为在这种情形下,在 xy 平面的点 (x_0, y_0) 的邻域将被映射成 $\xi\eta$ 平面上的一条曲线段,地理学家很难把这种图当作有用的图。因此,我们要考虑 $\Delta \neq 0$。在前面讨论中,我们已弄清这种仿射变换的性质,因此可以利用下述定理:点 (ξ_0, η_0) 的邻域可以在我们所关心的精确度以内从点 (x_0, y_0) 的邻域求出,即:使后者经过一个沿两垂直方向的纯粹变形,再转过一个适当的角度。你们会在蒂索的书中发现,他实际上对这个定理给出了一个特别清楚的推导。同时,你们可以在这里看到一个有趣的例子,说明关心应用的人如何努力满足对其所研究科目的数学要求。对于数学家来说,事情总是显得非常简单,但是了解一下这些应用的要求,仍然是有益的。

下面就来讨论点变换的一般情况。

§17.3 最一般的可逆单值连续点变换

我们至今讨论过的所有映射函数,都是连续和逐次可微的,事实上也都是解析的(可展为泰勒级数的)。但我们允许有多值,甚至无穷多值函数(例如对数)。现在提出我们的主要要求是:映射函数应该是无例外的单值可逆函数。我们也假设它们是连续的。但对导数的存在性等等不作任何假设。我们要问的是在这最一般的单值可逆连续变换下保持不变的几何图形性质。设想例如有一个橡胶做的曲面和固体,其上画有图形。如果橡胶物在不被撕破的前提下任意变形,那么这些图形中有哪些东西保持不变呢?

我们在处理这个问题时发现的所有性质,组成所谓拓扑学这样一个领域。我们可以称它为依赖于位置而完全不依赖于大小的那些性质的科学。这个名称来自黎曼。在 1857 年的著名论

文“解析函数理论”[1]中，他被函数论的兴趣吸引到这些研究之中。从那时起，拓扑学常常不在几何学书中叙述，而在函数论中用到它的地方加以讨论。但莫比乌斯却不这样，他在 1863 年写的论文[2]中是从纯粹几何意义去讨论拓扑学的。他称那些在可逆单值连续变形下互相变换的图形为“基本相关”图形，因为在那些变换下保持不变的性质是最简单的性质。

这里将以曲面的研究为限。我们应指出由莫比乌斯首先发现而被黎曼完全忽视的一个性质：界限，即关于一个曲面是单侧还是双侧的区分问题。我们已经在本卷第一章中讨论过单侧莫比乌斯带，通过其上的连续移动，能使人不知不觉地从一侧转到另一侧，因而两侧之间的区分失去任何意义。显然，这个性质在所有连续变形下都被保留着。因此，在拓扑中，我们从开始起就必须真正区分出单侧曲面与双侧曲面。

为了简单起见，这里只讨论双侧曲面，特别是因为在复函数论中通常只考虑它们。但单侧曲面的理论本质上也不难。对于在拓扑意义上的一个曲面，已发现有完全刻画它的两个自然数：它的边界曲线个数 μ，以及不能把它分成几部分的封闭切口个数 p 即所谓亏格。更准确说两个双侧曲面彼此能进行可逆单值连续变换（它们是“基本相关的”，现在称为同胚）的充分必要条件，是这两个曲面有相同的两个数 μ 和 p。这个定理证明起来就太啰嗦了，我只能用几个例子来说明这些数 μ 和 p。

我们设想 3 个曲面一个接一个摆在一起，一个是球面，一个是环面，再一个是双环面，如图 17.6 所示。每个都是封闭曲面，即没有边界曲线，因此 $\mu=0$，在第一个例子中，每个闭合切口把曲面分成两个

---

① *Journal für die recne und angewandte Mathematik*，第 54 卷，或《黎曼数学著作集》（1892 年，莱比锡，第 2 版），第 88 页。黎曼追随莱布尼兹，在这里把“分析”一词用于方法论上的本义，不是当作一个数学术语。

② “Theorie der elementaren Verwandtschaft，Berichte über die verhamdlungen der königlich Sächsischen Gesellschaft der Wissenschaften”，见 *mathematischphysikalische klasse*，第 15 卷，第 18 页，或《莫比乌斯著作集》，第 2 卷（莱比锡，1886 年），第 433 页。

分离的部分，故 $p=0$。在第二个例子中，经线 c 代表一个不把曲面分离成两部分的闭合切口。但在画了曲线 c 之后，任何其他闭合切口都会将曲面分成两部分，这正是我们所谓的 $p=1$ 的情况。在第三个例子中，两个分离的柄上有不同的经线 c_1 和 c_2，这表明 $p=2$。增加更多的柄，能得到具有任意数值 p 的曲面。另一方面，在这些曲面上钻若干个小孔或洞，每个小孔或洞增加一条边界，从而可以得到异于零的任意整数 μ。因此，我们实际上能做出具有任意数值 p 和 μ 的曲面，而且所有其他具有相同 p 和 μ 值的曲面必然和它们同胚，不论它们的外观如何不同。函数论中给出许多这种曲面的例子。

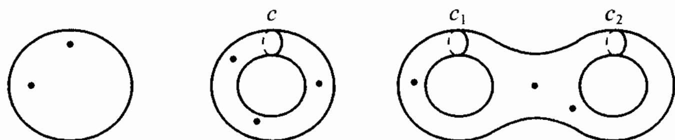


图 17.6

这里也必须解释一下黎曼引出的连通性这个术语。他用数 $2p+\mu$ 来表示，并称此曲面为 $2p+\mu$ 重连通的。如 $2p+\mu=1$，则曲面为单连通的，这时 $p=0$，$\mu=1$，即曲面与有一个孔的球面同胚，我们可以扩大那个洞，使球面连续变形成一个圆盘(图 17.7)。

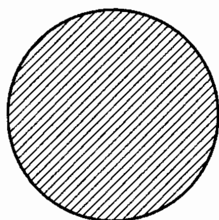


图 17.7

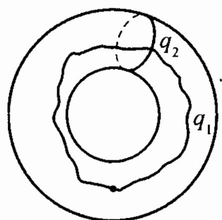


图 17.8

黎曼也引出了"交叉切口"的概念，即把一个边界点与另一个边界点相连的切口。因此，只有真正存在有边界的曲面时，即仅当 $\mu>0$ 时，才谈得上交叉切口。我们能证明下述定理：每个交叉切口把连

通性重数减少 1,特别是任何 $\mu>0$ 的曲面可用 $2p+\mu-1$ 个交叉切口变成单连通曲面。请考虑带有一个孔的环面(图 17.8)($p=\mu=1$),并从这个孔画出第一个交叉切口 q_1,它当然必定返回这个孔。然后我们从第一个切口出发,并以此切口为终点,画出第二个交叉切口 q_2,它完全与图 17.7 中环面闭合切口相似。于是连通性重数从$2\cdot 1+1=3$减为 1。

至于拓扑学的文献,在 M·德恩(M. Dehn)和 P·希加德(P. Heegard)所写,收于 *Enzyklopädie der mathematischen Wissenschaften*(ⅢAB3)的条文中,列出了一张详尽的单子,不仅包括曲面,而且包括任意扩充的流形。不过说实在的,所列文献都非常抽象。对于初学者来说,最好读比较好懂的东西,先用简单的例子来解释一般的思想,再进入抽象的理论①。

拓扑可应用于物理学,特别是位势理论。但它通过欧拉的多面体定理也渗入了中学的教学内容。对此,我想说几句话。欧拉发现,对任何具有 E 个顶点、K 条棱与 F 个面的普通多面体,必有关系式 $E+F=K+2$。如果我们使此多面体以任意方式作可逆单值连续变换(形),这些数字和关系将保持不变。故当 E, F, K 是球面或与它同态的曲面的任意划分的顶点、面和棱数时,只要所分成的各部分是单连通的,上述关系仍得以保持。我们立即可将此定理推广到有任意亏格的曲面。推广如下:如果把带有 P 个闭合切口而不会解体的曲面,用 K 条线段分成 F 个单连通部分,且设共得到 E 个顶点,则有 $E+F=K+2-2P$。我让你们去提出说明的例子,并思考定理的证明。当然这个定理还可以推广。

点变换的理论就讲到这里为止。下面把点转换成其他空间元素,以便对这类最重要的变换取得若干了解。

---

① 最新的著作是 B·冯·克里克贾多(B. V. Kerékjartó)的 *Vorlesungen über Topologie*(仅已出版一卷),柏林,施普林格出版社,1923 年。另一篇拓扑文章即将发表在 *Enzyklopädie der mathematischen Wissenschaften*,作者为 H·蒂采(H. Tietze)。

第十八章

空间元素改变而造成的变换

§18.1 对偶变换

最明显的情形是在二维区域中点与直线的交换，或在三维区域内点与平面的交换这种对应关系。我们只讨论前一种情形，并遵循普吕克于1831年在《解析几何论》第二部分所采用的思路。我们从解析叙述开始。

我们曾讨论过普吕克的第一个思想，是把直线方程

$$ux + vy = 1 \tag{1}$$

中的系数 u, v 放在与普通坐标同等的地位上，即视 u, v 为直线坐标，并以这两类坐标的类比"对偶"方法，建立起解析几何结构。因此，在平面上，曲线可作为由点方程 $f(x, y) = 0$ 给出的点的轨迹，也可作为由线方程 $g(u, v) = 0$ 确定的一阶无穷个直线族的包络，这两者是相互对应的。

只有除平面 E 外加上另一个平面 E'，且在 E 上的线坐标 u, v 与 E' 上的点坐标 x', y' 之间建立了对应关系之后，才能得到我们所要考虑的变换。因此，这类最一般的变换将由两个方程

$$u = \phi(x', y'), \ v = \chi(x', y') \tag{2}$$

给出，即每个 E' 上的点 (x', y')，对应于 E 上的将(2)式之值代入(1)式所得方程之直线。

(1) 首先让我们考虑这种变换的最简单的例子，即由方程

$$u = x', \ v = y' \tag{3}$$

给出的变换。据此变换，E'中的点(x', y')将对应于E中直线

$$x'x + y'y = 1。\tag{3a}$$

如果现在把平面E和E'叠在一起，使它们的坐标系一致，我们就会看到，这个方程代表点(x', y')相对于围绕原点的单位圆$(x^2 + y^2 = 1)$的极线。所以，我们的变换是大家知道的对于圆的极线关系（图 18.1）。

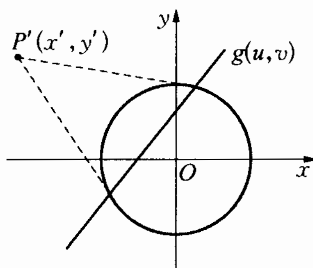


图 18.1

我们注意到，一个方程(3a)已足以代替两个方程(3)来确定这种关系，因为它是对应于任意点(x', y')的直线的方程。由于在这个方程中，x, y和x', y'是完全对称的，故平面E与E'在我们的关系中必然起到同样作用，即E中的每个点必然也对应于E'中的一条直线。当两平面放在一起后，我们把点看成在E或E'上是没有区别的。对第一个性质，我们称变换在狭义下是对偶的，而第二个性质为可逆性。因此不需要对这两个平面作任何区分，可以简单地说一个确定的极线与一个极点的对应，然后用“导出的流形”一章中所说的方法表达其互逆性质。

就此变换的其他性质而论，我只说明，E'内的点(x', y')的轨迹曲线，将对应于E内对应直线(u, v)的包络。

(2) 用类似于前面讨论最一般的“共线性”时所用的方法，不难证明，如果推广(3)式的假设，令u, v为x', y'的具有相同分母的线性分式函数

$$\begin{cases} u = \dfrac{a_1 x' + b_1 y' + c_1}{a_3 x' + b_3 y' + c_3}, \\[2ex] v = \dfrac{a_2 x' + b_2 y' + c_2}{a_3 x' + b_3 y' + c_3}, \end{cases} \tag{4}$$

就得到最一般的对偶关系。把 u, v 的这些值代入(1)式中,并乘以公分母,注意到9个系数 a_1, …, c_3 是任意的,就得到关于 x 与 y,以及 x' 与 y' 的最一般的线性方程

$$a_1xx' + b_1xy' + c_1x + a_2yx' + b_2yy' + c_2y - a_3x' - b_3y' - c_3 = 0。\tag{4a}$$

反之,每个 x, y 与 x', y' 的"双线性"方程,代表平面 E 与 E' 之间的一个对偶变换。因为如果假设一对坐标是常数,即设想在一个平面上有一个固定点,则该方程对其他两个坐标是线性的,并代表对应于此点的另一平面上的一条直线。

(3) 除非在方程(4a)的两个对称项有相同的系数,否则这个关系在前面定义的意义下一般是不可互逆的。在这种情况下,方程化为

$$Axx' + B(xy' + yx') + Cyy' + D(x + x') + E(y + y') + F = 0。\tag{5}$$

这样确定的变换是大家从圆锥曲线理论中已经了解的。它表达了方程为

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$$

的圆锥曲线的极线与极点的对应关系,每个这种极线关系是对偶的,并且是可逆的。

由此出发,我们可立即考虑更一般的,因空间元素改变而造成的一类变换,即相切变换。

§18.2 相切变换

如果我们从两平面的4个点坐标的一个任意高阶方程

$$\Omega(x, y; x', y') = 0 \tag{1}$$

出发，以代替双线性方程(4a)，就得到索费斯·李命名的相切变换。我们假设此方程满足连续性条件，按照普吕克的说法，它被称为准线方程。关于平面几何上的问题，在前述普吕克的著作中都可以找到有关研究①。首先使 x, y 固定，即考虑平面 E 中一定点 $P(x, y)$(图 18.2)。于是，对流动坐标 x', y' 而言，方程 $\Omega=0$ 代表平面 E' 中一条确定的曲线 C'，像刚才对待直线一样，使这条曲线作为空间 E' 的一个新元素对应于点 P，但当我们在 E' 上(例如在曲线 C' 上)取一固定点 $P'(x', y')$，则同一方程 $\Omega=0$ 在把 x', y' 当作固定坐标而把 x, y 当作流动坐标时代表平面 E 内一条曲线 C。当然此曲线 C 必然通过第一点 P。用这种方法，建立了平面 E 中点 P 与平面 E' 中 ∞^2 条曲线 C' 之间和平面 E' 中点 P' 与平面 E 中 ∞^2 条曲线 C 之间的对应关系，正如前面建立的点与直线间的对应关系一样。

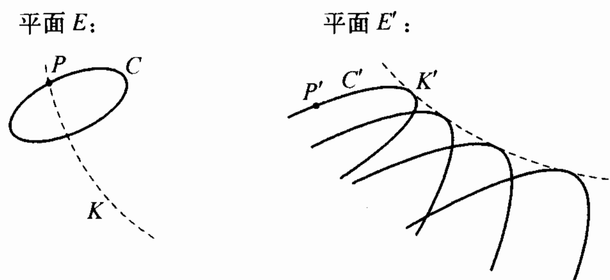


图 18.2

如果现在使平面 E 中点 P 沿一条任意曲线 K(用破折线表示)移动，则点 P 的每个位置将对应于平面 E' 中一条确定曲线 C'。为了从曲线 C' 所组成的单重无穷个族中求得平面 E' 中的一条能与平面 E 中曲线 K 相对应的曲线，我们对它应用在对偶关系中用过的包络原理：我们使按方程 $\Omega=0$ 确定的、在 E' 中与 K 的点对应的曲线

① 见普吕克的 *Neue Geometrie des Raumes gegründet aufdie Betrachtung der geraden Linieals Raumdement*，第 259—265 页。

族 C' 的包络 K' 对应于 K。显然，我们可以从 E' 中任意曲线 K' 出发作同样的讨论。于是从准线方程 $\Omega = 0$ 导出了两个平面之间的变换，使其一个平面上的每条曲线对应于另一个平面的确定曲线。

为了从解析的观点来理解这个讨论，让我们像在微分学中为了清楚起见所习惯做的那样，用具有短边的折线来代替曲线 K，并问什么东西对应于这样的一段边。当然，我们总会记住作为极限而过渡到曲线，所以实际上把折线的边理解为点 P 和它的移动方向（K 在点 P 的切线方向）。这一切组成所谓线段元素。现在在这个从 P 出发的方向上选一点 P_1（图 18.3），它具有坐标 $x+\mathrm{d}x$，$y+\mathrm{d}y$，其中 $\mathrm{d}x$，$\mathrm{d}y$ 是小的量并且最终接近于 0，但 $\frac{\mathrm{d}y}{\mathrm{d}x}$ 总有代表在点 P 给定方向的确定值 p。点 P 对应于 E' 中曲线 C'，其流动坐标 x'，y' 的方程为

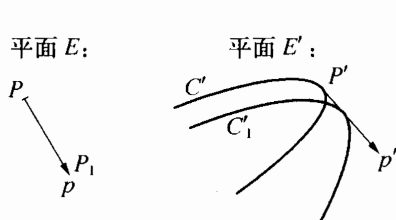


图 18.3

$$\Omega(x,\ y;\ x',\ y') = 0。$$

点 P_1 对应于曲线 C'_1，其方程为

$$\Omega(x+\mathrm{d}x,\ y+\mathrm{d}y;\ x',\ y') = 0。$$

按 $\mathrm{d}x$，$\mathrm{d}y$ 展开它，由于将过渡到极限，因此只保留线性项，我们得

$$\Omega(x,\ y;\ x',\ y') + \frac{\partial\Omega}{\partial x}\mathrm{d}x + \frac{\partial\Omega}{\partial y}\mathrm{d}y = 0。$$

这两个方程给出了曲线 C' 与曲线 C'_1 的交点坐标 x' 和 y'，过渡到极限即得曲线 C' 与包络 K' 的相切点。因为 $\frac{\mathrm{d}y}{\mathrm{d}x} = p$，我们可将这些方程写成

$$\begin{cases}\Omega(x,\ y;\ x',\ y') = 0,\\ \dfrac{\partial\Omega}{\partial x} + \dfrac{\partial\Omega}{\partial y}p = 0,\end{cases} \tag{2}$$

而且，在极限情况下，曲线 C' 与曲线 C_1' 在点 P' 有一个由方程 $\frac{\mathrm{d}y'}{\mathrm{d}x'}=p'$ 给出的公切线方向，它也是包络 K' 在 P' 的方向。因 $\Omega=0$ 是曲线 C' 的流动坐标 x' 与 y' 的方程，此切线方向由方程

$$\frac{\partial\Omega}{\partial x'}\mathrm{d}x'+\frac{\partial\Omega}{\partial y'}\mathrm{d}y'=0,$$

或

$$\frac{\partial\Omega}{\partial x'}+\frac{\partial\Omega}{\partial y'}p'=0 \tag{3}$$

确定。

因此，知道 K 上一点 P 及 P 处切线方向 p，那么对应于曲线 K' 上的点 P'，就同在 P' 的方向 p' 一起确定了。因此，通过我们的交换，在平面 E 的每个线元 x，y，p 与平面 E' 的一个确定的线元 x'，y'，p' 之间，用方程(2)和方程(3)建立了对应关系。

如果用这个方法来确定逼近对应曲线 K(或 K 的每个线元)的折线的每个边，那就在 E' 上得到逼近对应于曲线 K'(或 K' 的线元)的折线的边。因此，当令坐标 x，y 和斜率 p 在 K 上一切点所给出的值上移动时，即由解出 x'，y' 后的方程组(2)，给出曲线 K' 的解析表达式(图 18.4)。

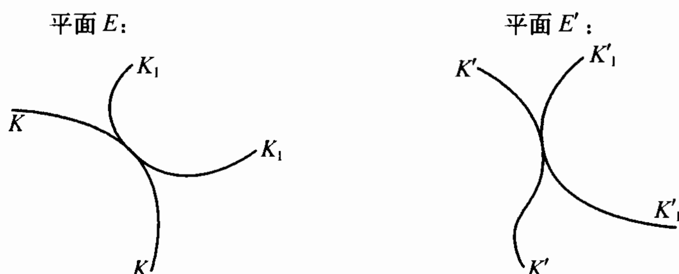


图 18.4

现在就清楚为什么李称这些变换为相切变换了。因为如在 E 上两曲线相切即意味着它们有一个公共直线元素；因此，在 E' 上

的对应两曲线必然有一个公共的直线元素，即有一个公共点和过此点的公共方向。这样，两曲线的相切性在此变换下是不变量，这就是此名称的含义。李也广泛地对空间研究了这些相切变换的理论。从 1896 年开始，他与 G・谢弗斯一起在名为 *Geometrie der Berührungstransformationen* 一书中作出了综合的表述，遗憾的是第 1 卷写完之后没有再写多少就中断了工作①。

简要地讨论了因空间元素改变而造成的变换理论之后，我将用若干具体例子来把这个理论说得更生动些，以便说明这些变换的应用。

§18.3 某些例子

我首先讲对偶变换及它们在代数曲线形式的理论中所起的作用。我们要问：如同关于圆锥曲线的互逆极关系那样，在对偶变换下典型曲线形状是如何变化的？当然，必须以若干典型情形为限。因此，首先考察具有奇数分支，且与每条直线相交于一个或 3 个实点的三次曲线。在下图(图 18.5)中有一条渐近线；但通过曲线射影变换，使得一条与它相交 3 个点的直线变到无穷，即可由此求得一个具有 3 条渐近线的曲线形状。在任何情况下，曲线均有 3 个实的拐点，而这些拐点都具有共线性的特殊性质。取此曲线的对偶，得到一条从任一点可作一条或 3 条切线的三类曲线，而拐点必然对应于一个尖点，仔细想一想就清楚了。在此导出的这种三类曲线(图 18.6)有 3 个尖点，而且这些尖点的切线必然经过一点 P'，它对应于 3 个拐点所在的直线 g。

现在对四次和四类曲线作类似的简短说明。一条四次曲线可能像一个凹进去的卵形线，事实上也存在着具有 2 个、3 个或 4 个凹进去部分的形式(图 18.7)，在第一种情况下，有两个实拐点和一条双

① 第 1 卷，莱比锡，1896 年。第 2 卷的前三章在李去世后发表在《数学年刊》第 59 卷(1904 年)上。

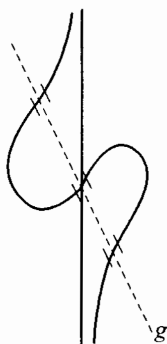


图 18.5

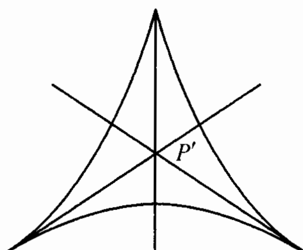


图 18.6

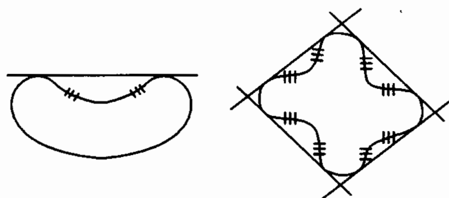


图 18.7

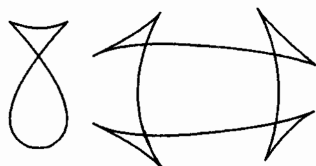


图 18.8

重切线；在其他情况下，则可能出现多达 8 个拐点和 4 条双重切线。如取其对偶，则必须补充说，一条二重切线的对偶是一个二重点。因此就会产生具有 2—8 个尖点和 1—4 个二重点的四类曲线，如图 18.8 所示。小心作出代数曲线图形，有特别令人陶醉之处。遗憾的是我在这里不能再详谈，只能满足于这几句简单的提示①。这些例子充分说明，初看起来似乎一点也不相像的东西，却因对偶变换而处于同样规律的支配之下。

现在讲讲相切变换理论的应用。十分有趣的是，和大多数真正

① 可看 F·克莱因《数学著作集》第 2 卷第 89 页及以后部分，第 136 页及以后部分，第 99 页及以后部分，柏林，施普林格出版社，1922 年。请看 *Über eine neue Art Riemannscher Flächen* 的两篇文章，以及 *Über den Verlauf der Abelschen Integrale bei den kurven 4 Grades* 的第一篇文章。

好的理论的想法一样，相切变换的想法也有广泛的应用领域。事实上，在建立这种理论之前很久，数学家已在利用相切变换。我现在特别想到的，是古老的齿轮原理。它构成机械运动学的专门的一章，是机械制造的关键。前面刚刚讲过的直线作图装置，也属于运动学的范围。我在这些讲座中经常讲的话，在这里同样也适用：我当然只能从每个数学分支中挑出一部分内容来讲，竭力通过一些简单的例子把这些内容的意义和重要性尽可能讲明白。在我的启发和鼓励下，我相信你们一定会从专门的著作中找到详细的内容来充实我的讲解。在整个运动学领域，A · 舍恩弗列斯(A. Schoenflies)为《百科全书》(IV_3)所写的文章是主要的指导文章，我建议你们一读。那篇文章还提供了有关大量文献的资料。

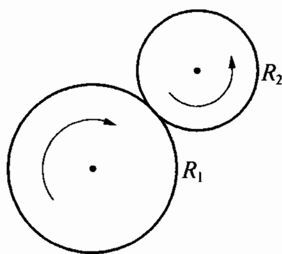


图 18.9

齿轮制造问题是把匀速运动从一个轮子传递到另一个轮子的问题，因为要同时传递力，所以让轮子一起滚动是不行的(图 18.9)。其中有一个轮子必须装上凸块(齿)，嵌入另一轮子的凹坑。因此，问题变成如何设计这些轮齿的形状，使得一个轮子的匀速转动带动另一个轮子的匀速转动，即使从几何角度来看，这也是一个十分有趣的问题。我将给出解答的最重要的部分：轮子之一可以根据齿与齿之间不能互相抵触等实用的限制加以任意选择，而第二个轮子的轮齿则随之完全被确定，因为事实上，它们是经过一个确定的相切变换而从第一个轮子的轮齿导出的。

我只简要地介绍一下这个定理是怎样得来的，不给出全部证明。首先注意到，我们只涉及两轮子的相互运动。因此，我们可以把其中之一 R_1 看成是固定的，而另一个轮子 R_2 除自身转动外，还绕 R_1 而转动。因此，R_2 中每一点在 R_1 的固定平面内描出一条旋轮线，且依点在 R_2 周线外部或内部而为长幅旋轮线(有尖点)或短幅旋轮线

(图 18.10)。由此推出，R_2 的运动平面的每一点对应于 R_1 平面的一条曲线。如果我们用已讨论过的办法，从表达这种对应关系的方程导出相切变换，那正好就是有关齿轮的相切变换。不难证明，在这个变换下相互对应的两曲线，实际上在这个运动中相互啮合。

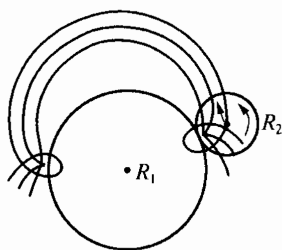


图 18.10

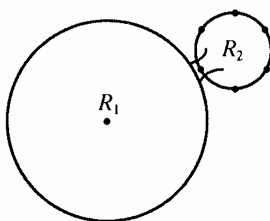


图 18.11

最后要说说上述理论原则在实际的齿轮制造中取何种形式。我只讲最简单的情形：驱动副齿轮的齿形。这里，R_2 的齿是简单的点(图 18.11)或小圆轴(因为点不能传递力)，即副齿轮。每个这样的小圆，在相切变换下，对应于一条与外摆线稍有不同的曲线，即与摆线平行而相距为副轮半径的曲线。当 R_2 转动时，这些圆在这些曲线上滚动，因而这些曲线必须竖在 R_1 上的齿侧面，以便 R_2 的圆弧齿正好咬合。在我向你们展示的模型中，这些曲线的开始部分可以作为 R_1 的齿形，每条曲线的宽度正好使轮齿咬合。

我再向你们介绍另外两个常用的轮齿形式：渐开线和摆线齿形①。在前一个形式中，两个轮子的齿形都是圆的渐开线(图 18.12)：即当一条线从一个圆周上拉开时生成的曲线，它的渐屈线是圆。在后一种形式中，轮齿由摆线的弧组成。

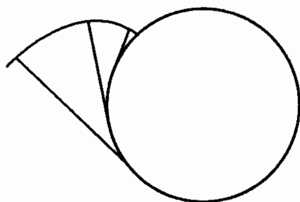


图 18.12

① 这些模型都是 F·谢林制造的(M·谢林公司，莱比锡)。

我希望至少已向你们介绍了有关空间元素改变而造成的变换的初步知识。在结束关于变换的第二大部分之前,我必须对我所讲的内容加以补充,讨论不能忽略的另一重要章节,即虚元素的应用。

第十九章
虚数理论

你们知道，虚数理论首先是在代数和分析里发展起来的，特别是在方程论和复变函数论中取得了最大的成就。但除此之外，更早的时候，数学家们已在解析几何中对变量 x 和 y 取了复值 $x = x_1 + ix_2$，$y = y_1 + iy_2$，因而除实数点外加上了一个大的复数点流形，但这种说法只是从分析里借用来的，没有对它指定任何适当的几何意义。

引入这个新概念的用处，当然在于：不必再由于实变量的限制而区分各种情况，并使一些定理得到统一的阐述而不致有例外的情况。我们在射影几何中也曾有过完全类似的考虑，结果引入了无穷远点和无穷远直线与无穷平面。我们所做的，恰当地说是在直观上可接受的空间的常态点之外添加所谓的"变态点"。

我们现在要同时进行两个推广。为此，我们像以前一样，引入齐次坐标。目前暂时考虑平面上的情形，我们令 $x : y : 1 = \xi : \eta : \tau$，并允许 ξ, η, τ 取复数值，将(0, 0, 0)除外。例如，我们考虑齐次二次方程

$$A\xi^2 + 2B\xi\eta + C\eta^2 + 2D\xi\tau + 2E\eta\tau + F\tau^2 = 0, \qquad (1)$$

并称满足它的所有数组(ξ, η, τ)(不论它们代表有限点或无穷远点)为一条二次曲线。有时也用圆锥曲线这个名字，但会造成误解。即使了解的人不至于误解，但不熟悉虚元素的人至少会如此。这种定义下的曲线不需要有一个实点。

我们现在联立方程(1)与线性方程

$$\alpha\xi + \beta\eta + \gamma\tau = 0。 \qquad (2)$$

把方程(2)当作一次曲线,即一条直线的定义。于是,这个方程组正好有两组公共值($\xi : \eta : \tau$),即一次曲线与二次曲线总有两交点,它们可以是实的或复的,有限的或无穷的,不同的或重合的。事实上,使这个定理产生例外的退化情形,是可以想象的。如果将方程(1)的左侧分解为两个线性因式,且其中之一与方程(2)相同,即如果二次曲线是一对直线,且方程(2)与其中之一相同,则方程(2)上的每个点都是公共点。这等于说,从所给方程组中消去一个变量而得到的二次方程的系数均化为零。当然,当给出方程组之一或两个的左端恒等于零($A = B = \cdots = F = 0$,或$\alpha = \beta = \gamma = 0$)时,会出现其他的退化情况。但我们将不考虑所有这些无关紧要的特殊情形,而考虑两条二次曲线。可以阐明这样的定理:两曲线总有 4 个公共点。

现在在空间引入齐次坐标 $x : y : z : 1 = \xi : \eta : \zeta : \tau$,并对它们指定除(0, 0, 0, 0)以外的任意复数值。这 4 个变量的线性齐次方程的全部解,称为一次曲面(平面);二次齐次方程的全部解,则为二次曲面。于是,如果抛去无关紧要的例外,一般情况下,一个二次曲面与一个平面的交线为一条二次曲线;两个二次曲面的交线为一条空间四次曲线,它与任一平面相交 4 个点。至于这些交线是否有实支,或是否完全在有限区域内,则都不确定。

彭赛列早在 1822 年就在他的《论图形的射影性质》一书中对圆和球面应用了这些概念。他并未用齐次坐标,也未用由齐次坐标才可能构成的精确公式,他是按对几何连续性的强烈直觉进行讨论的。为了准确了解他的著名结果,我们从圆的方程

$$(x-a)^2+(y-b)^2=r^2$$

出发,将之写成齐次坐标形式

$$(\xi-a\tau)^2+(\eta-b\tau)^2-r^2\tau^2=0。$$

它与无穷远直线 $\tau = 0$ 的交点,将由方程

$$\xi^2+\eta^2=0,\ \tau=0$$

给出，代表圆特征的常数 a，b，r 在此结果中并不出现。因此，每个圆与无穷远直线相交于相同的两个固定点

$$\xi : \eta = \pm \mathrm{i}, \ \tau = 0,$$

我们称之为虚圆点。用同样的方法可证明，每个球面与无穷远平面交于相同的虚圆锥曲线

$$\xi^2 + \eta^2 + \zeta^2 = 0, \ \tau = 0,$$

我们称之为虚球面圆。

其逆也是对的：每个二次曲线如果通过在它的平面上的虚圆点，则是一个圆；每个二次曲面如果包含虚球面圆，则是球面。这些就成了圆和球面的特征。

我们有意避免使用有时用到的"无穷远"圆点与"无穷远"球面等表述方式。事实上，从原点到虚圆点的距离，并不是像一下子会想到的那样，一定是无穷。反之，此距离具有形式 $\sqrt{x^2 + y^2} = \frac{\sqrt{\xi^2 + \eta^2}}{\tau} = \frac{0}{0}$，因此是不定式。根据趋向虚圆点的方式，可以对它指定任何极限值。类似地，任何有限点到虚圆点的距离是不定的，空间任意点到虚球面圆上一点的距离也如此。这是不足为奇的，因为我们要求这些虚圆点应与一个有限点的距离为 r（位于一个任意给定的半径为 r 的圆上），同时又应与它距离为无穷大。这个明显的矛盾，在解析式子里只能通过不定性才能调和。这些简单的道理必须搞清楚，特别是因为经常有些不正确的说法和写法。

有了虚圆点和虚球面圆，就有可能把圆和球的理论十分协调地包括在二次流形的一般理论里，而在初等的讨论中，就似乎存在某种差异。例如，在初等解析几何里，总习惯地说两个圆只有两个公共点，因为从方程中消去一个未知数后只产生一个二次方程。初等表示中没有考虑到两个圆在无穷远直线上还有两个公共的虚圆点。上

述的一般理论实际上为我们提供了 4 个交点，正好是两个二次曲线所需有的交点数。类似地，总是习惯地说两球面只相交于一个圆，而且可能是实的或虚的。但我们现在知道，各球面在无穷远平面上总是有公共的虚椭圆，加上有限圆，构成了一般定理所要求的两球面相交而得的四次曲线。

在这方面，我想就所谓虚变换说几句话。它表示一个具有虚系数的共线变换，并把我们考虑的虚点变成实点。因此，在虚圆点理论中，可以利用变换

$$\xi' = \xi,$$
$$\eta' = \mathrm{i}\eta,$$
$$\tau' = \tau。$$

这个变换把方程 $\xi^2 + \eta^2 = 0$ 变成方程 $\xi'^2 - \eta'^2 = 0$，并将虚圆点 $\xi : \eta = \pm \mathrm{i}$, $\tau = 0$ 变成实的无穷远点

$$\xi' : \eta' = \pm 1,\ \tau = 0。$$

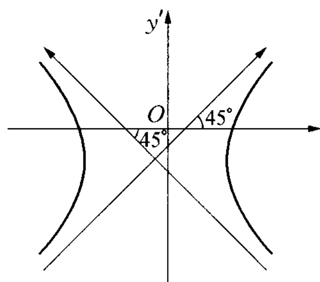


图 19.1

它们是与坐标轴成 45°角的两个方向上的无穷远点。因此，所有的圆被变换成通过这两个实无穷远点的圆锥曲线，即渐近线与坐标轴成 45°的等轴双曲线(图 19.1)。借助这些双曲线的图形，所有关于圆的定理都能得到解释。这对于某些研究，特别是对于空间的相应研究是十分有用的。我不准备超越本讲义的范围，只能限于这些简短的说明。比较完整的叙述，一般可以在射影几何教材中看到。

出现一个这样的问题：对这些虚点、平面、圆锥曲线等等，是否能采取纯几何的处理方法，而不像我们到目前为止所做的那样，只是从公式进行推导？彭赛列和斯坦纳等老一辈几何学家从未搞清楚这

一点。对斯坦纳来说，几何学中的虚数好像是幽灵，我们不能对它的存在获得一个清晰的观念，只有在冥冥之中才感到它存在的影响。只有冯·斯图特在前已提及的 *Geometrie der Lage*① 和 *Beiträge zur Geometrie der Lage*② 两书中才首先对这个问题作出了完整的答复。我们现在必须对他的见解作某些介绍。斯图特写的书是十分难读的，因为他的理论都是直接从最终形式推导出来的，不提解析公式，也没有关于归纳步骤的提示。人们能愉快地掌握的，只是随作者可能的思路发展而来的、供进一步思考的说法。斯图特写的两部著作，代表着他思想发展的两个不同的阶段，现在我要简单地加以说明。

他于 1846 年写的一本书主要是考虑实系数二次流形，我之所以说流形，是因为我想使维数不确定（直线、平面或空间）。例如请考虑在平面上的一条二次曲线，即一个实系数的 3 个变量的齐次二次方程

$$A\xi^2+2B\xi\eta+C\eta^2+2D\xi\tau+2E\eta\tau+F\tau^2=0。$$

对于解析的讨论来说，这个方程是否有实解，即此二次曲线有实分支或只有复数点，是无关紧要的事。对纯几何学家来说，问题是在后一种情况下应如何想象这样的曲线，如何用几何方法来确定它。在一维范围内，当我们用一条直线，例如 x 轴 $\eta=0$，与曲线相交时，也出现同样的问题。这个交点不论是不是实的，都由实系数方程

$$A\xi^2+2D\xi\tau+F\tau^2=0$$

给出。而问题在于出现复根的情况下，是否能对它们赋予几何意义？

斯图特的思想如下：首先，他不考虑二次曲线，而考虑我们曾讨论过的二次曲线的极线系统，即由方程

① 纽伦堡，1846 年。
② 纽伦堡，1856—1860 年。

$$
\begin{aligned}
A\xi\xi' + B(\xi\eta' + \xi'\eta) + C\eta\eta' + D(\xi\tau' + \xi'\tau) \\
+ E(\eta\tau' + \eta'\tau) + F\tau\tau' = 0
\end{aligned}
$$

给出的一个对偶互逆关系。由于系数是实的,因此这是一个完全实的关系,它给出了每个实点与实线的对应关系而不论曲线本身是否是实的。另一方面,这个极线系统把曲线作为本身在其极线上那些点的总体而完全地确定下来。至于这些点是否实际存在的问题,则放在一边。但在任何情况下,极线系统总是由上述方程确定的二次曲线的实表示,且能代替曲线本身,成为研究的对象。

如果现在用 x 轴与这条曲线相交,即令 η 与 η' 等于 0,则用类比的方法,我们将得到一个由方程

$$
A\xi\xi' + D(\xi\tau' + \xi'\tau) + F\tau\tau' = 0
$$

给出的一维实极线系统,总是使两个实点处于彼此可互逆的关系中。x 轴与曲线的各个交点是在这个极线系统中的两个自相对应的点,称之为基本点或阶点。它们可以是实的或虚的,但它们只有次要意义,主要的还是在于这个极线关系是它们的实表示。

为了指定在这个一维极线关系中相互对应的两个点$\left(\frac{\xi}{\tau}, \frac{\xi'}{\tau'}\right)$,我们应用对合的点对这个术语(它起源于 17 世纪的迪萨格斯(Desargues)),并按基本点是实或虚以及它们相重合的转移情形而将对合分成两个主要类型。在这里,对我们而言主要问题是对合概念本身;至于在各个情况下的区别,即关于二次方程根的性质的问题,是第二位的。

这些讨论很容易推广到三维空间,实际上并未对虚数提供一个解释,但就二次流形来说,仍然为区别实与虚提供了一个基准点。每个二次流形由一个实的极面系统表示,而且能像对流形的实方程作解析运算一样,对这个极面系统作几何运算。

举一个能充分说明这一点的例子。考虑一条二次曲线,即在平面上给出的一个极线系统,还考虑一条直线。根据此曲线是否有实点以

及如有实点,此直线与它是否相交于实点而直观地给出了许多可能的情况。在任何情况下,平面的极线系统将在直线 g 上(图 19.2)建立一个线性极线系统,即一个对合。g 上的每一点 P,在第一个系统里对应于一条极线 p',它与直线 g 相交于点 P'。点 P, P'取遍了有关对合。我们也可以问及基本点问题,并确定它们是实或虚。由这个讨论开始出现的方程所推导出的结果,都用这一切说法变成了几何语言。

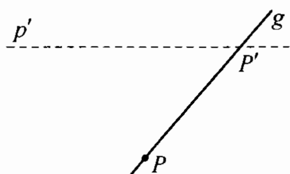


图 19.2

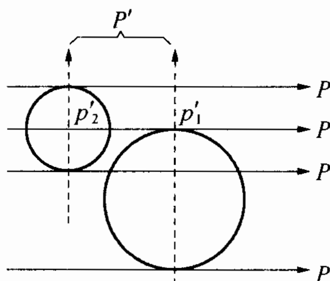


图 19.3

我们将把这些考虑特别地运用于虚圆点和虚球面圆。我们以前说过,任何两圆与无穷远直线相交于同样两个点,即虚圆点。从几何上来说,这意味着,它们的极线系统在无穷远直线上建立了同样的一维极线系统,即同样的对合。事实上,如果我们画几条从无穷远点 P 到圆的切线,那么,作为这些切线的切点的连线,极线 p_1'将垂直于它们的公共方向(图 19.3)。因为通过同样无穷远点的所有直线是平行的,所以相对于第二个圆,点 P 的极线 p_2'将垂直于 p_1'所垂直的方向,因而与 p_1'平行。换句话说,p_1'与 p_2'与无穷远直线相交于同一点 P'。因此,在同一极线系统中所有圆的极线系统与无穷远直线相交于一点,即所谓"绝对对合",它的点对从任何有限点去看,都以相互垂直的方向出现。

现在把这些思想用解析语言来表达。如果从圆的齐次方程

$$(\xi-a\tau)^2+(\eta-b\tau)^2-r^2\tau^2=0,$$

或

$$\xi^2+\eta^2-2a\xi\tau-2b\eta\tau+(a^2+b^2-r^2)\tau^2=0$$

出发，则对应的极关系是

$$\xi\xi'+\eta\eta'-a(\xi\tau'+\xi'\tau)-b(\eta\tau'+\eta'\tau)$$
$$+(a^2+b^2-r^2)\tau\tau'=0。$$

如果令 $\tau=\tau'=0$，则可由此得出在无穷远直线上的关系

$$\xi\xi'+\eta\eta'=0,\ \tau=0,\ \tau'=0。$$

这些方程事实上与原来圆的常数 a，b，c，r 无关。进而，根据第一方程，由解析几何推出，从点$(\xi,\ \eta,\ 0)$与$(\xi',\ \eta',\ 0)$作出的两条直线是互相垂直的，所以实际得到了上面的定理。

对空间球面也有完全类似的结果。它们在无穷远平面上都产生同样的，由方程组

$$\xi\xi'+\eta\eta'+\zeta\zeta'=0,\ \tau=0,\ \tau'=0$$

给出的所谓绝对极线关系。因为第一个方程表明方向 $\xi:\eta:\zeta$ 与 $\xi':\eta':\zeta'$ 是相互垂直的，所以，每个无穷远点 P，对应于在垂直于从一个有限点到 P 的方向的平面上的无穷远直线。故我们得到了关于虚球圆定理的一个实际几何等价物。

可能会说，在这个讨论中不如把虚数抛去，不加解释。首先对单独的虚点、直线、平面作出实际解释的，是斯图特。他在 1856—1860 年所写的 *Beiträge* 中，通过本定理的推广而给出了这种解释。下面我也要介绍一下这个解释，因为它实际上是简单而又巧妙的，其所以感觉奇怪与难懂，仅在于斯图特的抽象表示方法。我将采用斯托兹(Stolz)在 1871 年给出的解析方法①。斯托兹和我那时一起都在哥

① “Die geometrische Bedeutung der complexen Elemente in der analytischen Geometrie”,《数学年刊》第 4 卷，第 416 页，1871 年。

廷根工作，我以前从来不敢读斯图特的书，他却先我而读了，因此在与他个人往来中，我不仅了解了斯图特的这些思想，而且还了解了其他许多有趣的思想，后来我自己对斯图特的思想进行了大量的研究。这里，我只希望给出斯图特的最重要的思路而不讲其细节，仅以平面为限。

首先设用复坐标(ξ, η, τ)给出一个虚点，并将其实部和虚部分开如下

$$\xi = \xi_1 + \mathrm{i}\xi_2,\ \eta = \eta_1 + \mathrm{i}\eta_2,\ \tau = \tau_1 + \mathrm{i}\tau_2。\tag{3}$$

现在我们希望构造一个实的图形，借以解释点 P，并使其"关系成为可投影的"，较准确地说，使它在任何实投影变换下保持不变。

(1) 第一个必要步骤是把注意力集中在由点 P 坐标的实部和虚部分别组成的点 P_1，P_2

$$P_1:\ \xi_1,\ \eta_1,\ \tau_1;\ P_2:\ \xi_2,\ \eta_2,\ \tau_2。\tag{3a}$$

这两点是不同的，即不可能有关系 $\xi_1 : \eta_1 : \tau_1 = \xi_2 : \eta_2 : \tau_2$，否则 $\xi : \eta : \tau$ 的性质将和 3 个实数的一样，因而代表一个实点。因此，P_1，P_2 确定一个方程为

$$\begin{vmatrix} \xi & \eta & \tau \\ \xi_1 & \eta_1 & \tau_1 \\ \xi_2 & \eta_2 & \tau_2 \end{vmatrix} = 0 \tag{4}$$

的实直线 g。这条直线既包含点 P，也包含其共轭虚点 $\overline{P}$，其坐标为

$$\bar{\xi} = \xi_1 - \mathrm{i}\xi_2,\ \bar{\eta} = \eta_1 - \mathrm{i}\eta_2,\ \bar{\tau} = \tau_1 - \mathrm{i}\tau_2。\tag{3b}$$

因为 P 与 $\overline{P}$ 的坐标均满足直线方程(2)。

(2) 当然，这样作出的点对 P_1，P_2 绝不能作为虚点 P 的代表，因为它们本质上依赖于 ξ, η, τ 的各个值，而点 P 只与这些值的比有关。如果不用 ξ, η, τ，而用 ξ, η, τ 分别与一个任意复常数 $\rho = \rho_1 +$

$\mathrm{i}\rho_2$ 之积来代替，写成

$$\begin{cases}\rho\xi=\rho_1\xi_1-\rho_2\xi_2+\mathrm{i}(\rho_2\xi_1+\rho_1\xi_2),\\ \rho\eta=\rho_1\eta_1-\rho_2\eta_2+\mathrm{i}(\rho_2\eta_1+\rho_1\eta_2),\\ \rho\tau=\rho_1\tau_1-\rho_2\tau_2+\mathrm{i}(\rho_2\tau_1+\rho_1\tau_2),\end{cases}\tag{5}$$

那么同样的点 P 就得到了表示。但如果将其实部同虚部分出，我们得到代替点 P_1, P_2 的另外两个实点，其坐标为

$$\begin{cases}P_1'：\xi_1'：\eta_1'：\tau_1'=(\rho_1\xi_1-\rho_2\xi_2):(\rho_1\eta_1-\rho_2\eta_2):(\rho_1\tau_1-\rho_2\tau_2),\\ P_2'：\xi_2'：\eta_2'：\tau_2'=(\rho_2\xi_1+\rho_1\xi_2):(\rho_2\eta_1+\rho_1\eta_2):(\rho_2\tau_1+\rho_1\tau_2)。\end{cases}\tag{5a}$$

如果我们考虑由所有值 ρ_1, ρ_2 给出的全部点对 P_1' 与 P_2'，则我们有一个仅由比 $\xi:\eta:\tau$ 所求出的几何流形，即能拿来代表 P 的“几何”点 P，而且与 P 的联系实际上是投影联系。因为如果用任何实线性方法变换 ξ, η, τ，则 ξ_1', η_1', τ_1' 与 ξ_2', η_2', τ_2' 会得到同样的代换。

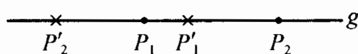


图 19.4

(3) 现在，为了研究这些点对总体的几何性质，我们首先注意到，不论 ρ 的值如何，点 P_1' 与 P_2' 都在直线 P_1P_2 上(图 19.4)，因为它们的坐标显然满足方程(4)。而且，如果让 ρ 取遍所有复数值，即 ρ_1 与 ρ_2 取遍所有实数值(一个公共实因子不会造成实质性差异)，则 P_1' 取遍 g 的所有实点；而 P_2' 总是代表在 g 上唯一与 P_1' 对应的第二个实点。故对 $\rho_1=1$, $\rho_2=0$，作为对应点而得到 P_1 与 P_2。如果引入比例

$$\frac{\rho_2}{\rho_1}=-\lambda,$$

则对应关系会显得更为清楚。于是我们有

$$对\ P_1'：\xi_1'：\eta_1'：\tau_1'=(\xi_1+\lambda\xi_2):(\eta_1+\lambda\eta_2):(\tau_1+\lambda\tau_2),$$

对 P_2'：$\xi_2' : \eta_2' : \tau_2' = \left(\xi_1 - \frac{1}{\lambda}\xi_2\right) : \left(\eta_1 - \frac{1}{\lambda}\eta_2\right) : \left(\tau_1 - \frac{1}{\lambda}\tau_2\right)$。

(5b)

(4) 从这些公式中也能推知，当 λ 变化时，P_1' 与 P_2' 变成为 g 上一个对合的所有点对。因为如果在 g 上引入一个一维坐标系，则点 P_1' 与 P_2' 的齐次坐标分别变成方程(5b)中参数 $\lambda_1' = \lambda$，$\lambda_2' = -\frac{1}{\lambda}$ 的线性整函数。因此，两参数之间的方程 $\lambda_1' \cdot \lambda_2' = -1$ 在 P_1' 与 P_2' 的线性坐标之间产生一个对称双线性关系。

(5) 对合的基本点，即由 $\lambda = -\frac{1}{\lambda}$ 或 $\lambda = \pm i$ 给出的相互对应点，它们都是虚的，其中之一是我们的出发点 P，另一是共轭虚点 $\overline{P}$。到目前为止，我们只是对斯图特的理论作出了一个新的解释。除点 P 外，我们也考虑了点 $\overline{P}$，它和 P 一起形成一个由实二次方程确定的二次一维流形，于是我们构造出了对合作为它的实代表。我提醒你们注意，如果我们知道点对中的两个，例如 P_1，P_2 与 P_1'，P_2'，这样的对合就确定了。如果这个对合有虚的基本点，其必要充分条件是这些点对相互交错，即点 P_1' 与 P_2' 中有一个应在点 P_1 与 P_2 之间，而另一个则在它们之外。

(6) 为了完全解决我们的问题，我们只需用一个方法把 P 与 $\overline{P}$ 的公共代表变换成只是 P(或只是 $\overline{P}$)的一个代表。斯图特于 1856 年发现了这样一个方法，这是他的光辉思想的结晶。如果 λ 从 0 到 $+\infty$ 然后返回经负值到零取遍所有实数值，则具有坐标 $\xi_1 + \lambda\xi_2 : \eta_1 + \lambda\eta_2 : \tau_1 + \lambda\tau_2$ 的点 P_1' 将以一个完全确定的方向走遍直线 g(图 19.5)。不难证明，如果从点 P 的坐标乘以一个任意数 P 出发，即考虑点 $\xi_1' + \lambda\xi_2'$，…，则将在 g 上导出完全同样的方向。而且，在 P 的实射影变换下，像点的箭头方向将作为同样的变换结果而从刚刚确定的方向导出。于是，如果使这个箭头方向对应于原来点 $P(\xi_1 + i\xi_2, \cdots)$ 的方向，就能满足我们的要求。由于共轭虚点 $\overline{P}$ 的坐标为

$\xi_1+\mathrm{i}(-\xi_2)$，…，所以必须相应地指定正的增加的 λ 为 P 的移动方向，与刚才对直线 g 确定的方向相反。由此得到所需要的区分：我们用实数 λ 的正和负的进向来区分 $+\mathrm{i}$ 与 $-\mathrm{i}$。

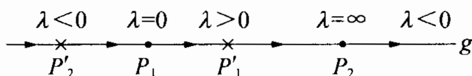


图 19.5

这样，为了表示虚点 $\xi_1+\mathrm{i}\xi_2$，$\eta_1+\mathrm{i}\eta_2$，$\tau_1+\mathrm{i}\tau_2$，我们至少有了下述法则，以建造一个唯一的、射影不变的实几何图形：作出点 $P_1(\xi_1:\eta_1:\tau_1)$ 与 $P_2(\xi_2:\eta_2:\tau_2)$，以及它们的连线 g，及在 g 上的点对合(或在 g 上的另外点对)，其中点

$$P_1'(\xi_1+\lambda\xi_2:\eta_1+\lambda\eta_2:\tau_1+\lambda\tau_2)$$

与
$$P_2'\left(\xi_1-\frac{1}{\lambda}\xi_2:\eta_1-\frac{1}{\lambda}\eta_2:\tau_1-\frac{1}{\lambda}\tau_2\right)$$

总是成对的。最后，按 λ 的正增量而得到 P_1' 的移动方向，我们加上箭头。

(7) 剩下还需要证明的，是反过来的问题：每一条加有定向箭头的直线，以及其上两个相互交错的点对 P_1，P_2 与 P_1'，P_2'(或一个没有实点对的对合)构成的实图形，代表一个且仅一个虚点。我不需要对此作详细的证明。但选择一个合适的实常数因子以后，不难对 P_2 的坐标给出这样的值 ξ_2，η_2，τ_2，使得 P_1' 和 P_2' 的坐标和

$$\xi_1+\lambda\xi_2:\eta_1+\lambda\eta_2:\tau_1+\lambda\tau_2 \text{ 与 } \xi_1-\frac{\xi_2}{\lambda}:\eta_1-\frac{\eta_2}{\lambda}:\tau_1-\frac{\tau_2}{\lambda}$$

成比例，或与之相同，即：使得所设对合范围的点对具有坐标 $\xi_1\pm\mathrm{i}\xi_2$，… 至今仍属任选的 λ 的符号，应作这样的选择：当 λ 从 0 向正数增加时，点 $\xi_1+\lambda\xi_2:\eta_1+\lambda\eta_2:\tau_1+\lambda\tau_2$ 的移动方向与原来的箭头

方向一致，于是，根据前面的观点，具有坐标 $\xi_1+\lambda\xi_2$，… 的点 P，将实际代表具有给定箭头方向的给定的对合。而且可以证明，如果从对合的其他点对出发，会得到同样的坐标比例，即得到同样的点 P。

讨论了点的有关问题之后，我们可以用对偶原理来解决平面上的直线问题。于是，我们可以用一个实点（或一个没有重实射线的线束对合），加上线束内的一个确定的旋转方向，对复直线作出实的唯一的表示。

这些结果，也允许借助实的几何图像的有形性质来表示复与实元素之间的所有关系。这也是这些结果的实际价值。为了用一个具体的例子来说明这一点，我将向你们说明“一个点（实或虚）在一条直线 g（实或虚）上”这句话在这种表示中的含义。当然，这里有 4 种情况：

(1) 实点与实线。

(2) 实点与虚线。

(3) 虚点与实线。

(4) 虚点与虚线。

情况(1)不需要做什么解释，它组成通常几何学中的一个基本关系。在情况(2)中，所给实点也必须在共轭虚线上，因此它必须与我们用来代表虚线的线束的顶点一致。类似地，在情况(3)中，该实直线必须与用来表示所给虚点的点对合的直线一致（图 19.6）。情况(4)是最有意思的（图 19.7）。在这种情况下，显然，共轭虚点必须在

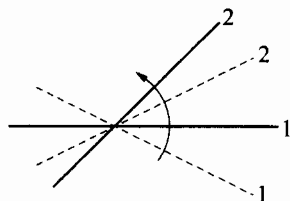


图 19.6

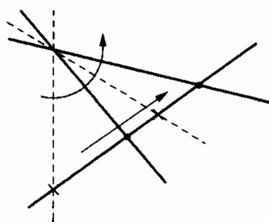


图 19.7

共轭虚直线上。由此推得，表示 P 的对合直线的每一个点对必须在表示 g 的对合线束的一对直线上，即这两个对合必须是相互投影的，而且它们的箭头也是相互投影的。

综合这个讨论，我们可以说，如果把带有给出方向箭头的全部给定对合图形作为新元素增加到平面上全部实点和实直线中去，就可以得到一个把复元素考虑在内的、解析几何的平面的完全实的图形。这里，也许只要大致说一说怎样做出复几何的这个实图形。我将沿着前面一些初等几何定理的一般提出顺序进行。

(1) 我们从存在定理出发，它把普通几何扩充范围以内出现的元素准确地考虑进去了。

(2) 然后考虑连接定理，即在第一点确定的扩充范围内，过两点有一条直线且仅有一条直线，两条直线有且仅有一公共点。这里正和上面一样，要根据给出元素的现实分成 4 种情况。有意义的是确定在什么点与线对合中可以得到这些复关系的像。

(3) 关于顺序的定律，和实关系相比较，这里出现了一个全新的情况。特别是，在一条直线上的所有实点与复点组成了一个二维连续统，过一固定点的所有直线也如此。事实上，每一个学过复变函数论的人，都习惯于用平面的所有点来表示复变数的集合。

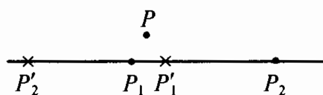


图 19.8

(4) 关于连续性定理，我只指出，如何表示任意靠近一个实点的各个复点。为此，通过实点 P(或通过一个邻近的实点)画一条线，并在其中取两对互相交错的点 P_1，P_2 与 P_1'，P_2'(图 19.8)，使不同对的两个点 P_1，P_1'位于靠近点 P 之处，如果使 P_1 与 P_1'移动至重合，则由这些点对决定的对合将退化，即两个复的二重点与 $P_1=P_1'$重合。于是，由对合(具有某个箭头)所表示的两个虚点，各连续地移动到一个靠近 P 的点或移动到点 P 本身。当然，必须小心地引入连续性表示，以便有效地使用它们。

这样的整个构造与普通的实几何学比较是啰嗦的，但另一方面，它又包含着无比丰富的内容。特别是，它把代数流形完全弄清楚了，从几何观点把它看成是其中实与复元素的总体。因此，两个 m 与 n 次曲线一般有 $m \cdot n$ 个公共点这类代数基本定理或贝佐(Bézout)定理，就可以用几何图形明显地表示出来。为了实现这个目标我们应更小心地去推导定理。而这类研究所需要的一切关键资料，都可以在有关文献中找到。

在大多数情况下，尽管这种几何解释在理论上有它的好处，但确实会带来一些复杂的问题，使我们只好限于讨论它的理论可能性，而实际上却回到了比较简单的基点：一个复点是一组复数坐标值，在一定范围内，能用它作为实点进行运算。事实上，在完全不考虑所有理论问题的情况下，虚元素的这种用法在处理虚圆点与虚球面圆中已始终证明是有效的。正如我们所了解的，彭赛列是第一个这样用的人，法国其他的著名几何学家也跟着使用。在德国，特别是数学家李，曾使用这种虚数的概念，并取得极大的成功。

虚数的讨论就到这里为止，我准备这样结束第五大部分，并转入新的内容。

第六部分

几何及其基础的系统讨论

第二十章

系统的讨论

在这一章里，我们要用几何变换来阐明整个几何领域的分类，即用一个观点来纵观各个部分及其相互关系。

§20.1 几何结构概述

这里所涉及的想法，就是我在 1872 年爱尔朗根大纲[①]中系统地加以研究的内容。至于后来这些想法的发展，可以参看 G·法诺(G. Fano)在《百科全书》(Ⅲ A. B. 4b)中写的一篇文章："Die Gruppentheorie als geometrische Einteilungsprinzip"。

(1) 就像以前的做法一样，我们将前后一贯地运用分析来阐明几何关系，把空间上点的全体用 3 个坐标 x, y, z 值的全体来代表。因而空间上的每一个变换，对应于这些坐标的某一个变换。我们讨论一开始就已知的 4 种特别重要的变换，它们由 x, y, z 的某些线性代换来表示，即平移、环绕原点 O 的旋转、关于 O 的反射、以 O 为中心的相似变换。

(2) 可能认为，引入坐标会使 3 个独立变量(x, y, z)的分析和几何完全等同起来。但至少在特定的意义上并非如此。正如我已强调的，几何所研究的只是在前一点所提到的线性代换下仍保持不变

---

① "Vergleichende Betrachlungen über neuere geometrische Forschungen"，爱尔朗根，1872 年。重印于《数学年刊》，第 43 卷，第 63 页及以后部分，1893 年。也可看 F·克莱因《数学著作集》，第 1 卷，第 460 页及以后部分，柏林施普林格出版社，1921 年。

的那样一些坐标之间的关系,不论这些线性代换被看做是坐标系的变化也好,看做空间变换也好。因此,几何是那样一些线性代换的不变量理论。另一方面,坐标之间的一切非不变方程,例如某一点有坐标(2, 5, 3)之类说法,意味着参照一旦固定就永远固定的确定的坐标系。诸如此类的讨论本来属于逐点研究并分别考虑各点性质的学科,即属于地形测量学的范围,或也可以说属于地理学的范围。为了有助于理解,我请你们注意有关几何性质的几个例子。当说到两点相隔一距离时,意味着一旦长度单位选定之后,就可以根据其坐标(x_1, y_1, z_1)和(x_2, y_2, z_2)列出式子

$$\sqrt{(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2}。$$

这个式子在上述一切线性代换下或是保持不变的,或只是被乘上一个与点的特定位置无关的因子。又如,谈到两条直线的交角、圆锥曲线的主轴和焦点,都必须类似地来理解。

这些几何性质的全体,我们称为度量几何,以区别于其他种类的几何。至于其他种类的几何,则根据一定的原则把度量几何中某些定理提出来加以考虑,就可以得出。因此,所有这些新的几何科目至少从直接目的来说是范围最广的度量几何的组成部分。

(3) 我们先来看已经仔细研究过的仿射变换,即x, y, z的整线性代换:

$$\begin{cases} x' = a_1x+b_1y+c_1z+d_1, \\ y' = a_2x+b_2y+c_2z+d_2, \\ z' = a_3x+b_3y+c_3z+d_3。\end{cases}$$

在这种变换下,第一点中提及的一切变换都被当作特例包括进去了。从全部几何概念、几何定理中,我们选取了在一切仿射变换下都保持不变的那个较小的部分。这些概念及定理的全体,我们看做是几何学的第一个新的部分,即所谓仿射几何或仿射变换的不变式理论。

根据我们对于仿射变换的知识,可以立刻选出这种几何的概念

及定理。这里只提几点：在仿射几何中不能讨论距离和角。圆锥曲线主轴的概念以及圆和椭圆的区别，同样消失。但是仍保留着有限及无限空间的区别以及由这种区别而来的一切概念，如两条直线平行的概念，圆锥曲线之区分为椭圆、双曲线、抛物线，等等。此外还保留着圆锥曲线的中心及直径概念，特别是共轭直径的关系。

(4) 下面进而讨论投影变换，即引入线性分式变换：

$$x' = (a_1x + b_1y + c_1z + d_1) : (a_4x + b_4y + c_4z + d_4),$$
$$y' = (a_2x + b_2y + c_2z + d_2) : (a_4x + b_4y + c_4z + d_4),$$
$$z' = (a_3x + b_3y + c_3z + d_3) : (a_4x + b_4y + c_4z + d_4)。$$

它将仿射变换作为特例包括在内。在这些变换下保持不变的几何性质，当然也必须属于仿射几何的范围。因而从仿射几何中作为投影变换的不变式理论分出所谓投影几何。这种逐步筛选的办法，即从度量几何中分出仿射及射影几何的过程，可以比之于化学家用越来越强的试剂从化合物中分离出更加贵重成分的过程。我们所分出的，首先是仿射变换；其次是投影变换。

至于投影几何的定理，应该强调，在仿射几何中无穷大的特殊作用及与无穷大有关的概念，现在全没有了，只有一种常态的二次曲线。但是举例来说，还剩下极点和极线的关系，同样还剩下用投影线来生成二次曲线的问题，这一点我们前面已讨论过(见投影变换部分)。

利用同样的原则，下面我们也可以从度量几何中导出其他几何类别，其中最重要的一个类别是反演几何。

(5) 反演几何，这包括一切反演变换下仍然成立的度量几何定理之总和。在这种几何中，直线和平面的概念没有独立的意义，它们相应地作为圆或球面概念的特例而出现。

(6) 最后让我提出另一种几何类别，它在某种意义上是通过最仔细的筛选而得出的，因而包括的定理极少。这就是我们前面提及的拓扑学，它讨论了在所有单值可逆连续变换下的所有不变性。

为了避免在所有这种变换下无穷大变成本身因而得到一个特殊地位，我们或者将种种投影变换结合起来，或者将种种反演变换结合起来。

下面我们将引入群的概念，使上述划分方案的定义更加清晰。我们已经知道，如果变换集中任意两个变换的复合仍是集中的一个变换，每一个变换的逆变换仍属于此变换集，则称这样的变换集为群。群的例子有：一切运动的全体或一切射影变换之全体，因为两个运动复合成为一个运动，两个射影变换复合成一个射影变换，在两种情况下，每个变换都存在着一个逆变换。

如果回顾我们所划分的各种几何，我们可以看到，在各种几何中起划分作用的变换，总是形成为群。首先，使度量几何的种种关系保持不变的一切线性变换(平移、旋转、反射、相似变换)，显然形成为群，称为空间变换的主群。对于仿射几何中仿射变换的仿射群，也很容易确定其类似的意义；对于投影几何中一切投影变换组成的投影群也是如此。在把任何反演变换同主群里的变换结合起来而得到的一切变换下，反演几何的定理也是仍然成立的。这一切又形成为群，即反演变换群。最后对于拓扑来说，它必须涉及一切连续单值可逆变换群。

现在我们想确定，各个群中的一个单独运算依赖于多少个独立的参数。在主群中，运动涉及 6 个参数还须加上单位长度变化的参数，所以一共有 7 个参数。我们用 G_7 来表示主群。一般仿射变换方程含有 $3 \cdot 4 = 12$ 个任意系数；投影变换含有 $4 \cdot 4 = 16$ 个任意系数，但其中一个公因数是非本质的。因此仿射变换群为 G_{12}，投影变换群为 G_{15}，反演变换群为 G_{10}。最后，一切连续变换群则不具有有限个参数，其运算取决于任意的函数，如愿意的话，说取决于无穷多个参数也可以，不妨说这一群是 G_∞。

在刚才讨论的各类几何和变换群之间的联系中，有一个基本原则贯穿其间，可以作为一切可能建立的几何特征。我编的爱尔朗根

数学纲领中的主导思想就是如此：给出任一把主群作为子群包含在内的空间变换群，从该群的不变式理论建立某一类几何。每一类几何都可以用这个方法建立。因此各类几何都以其有关群作为其特征，相应的群就成了我们所考虑的主要对象。

在我编的教学大纲中，只有我概括的前 3 种情况，彻底贯彻了这个原则。这些情况是最重要或是大家最熟悉的，我们要花点时间来加以讨论，并特别注意其中的过渡情况。

我们采取与刚才相反的顺序，从射影几何开始，即从一切投影变换所形成的 G_{15} 出发，不妨将其记为齐次形式

$$\begin{cases}\rho'\xi' = a_1\xi + b_1\eta + c_1\zeta + d_1\tau, \\ \rho'\eta' = a_2\xi + b_2\eta + c_2\zeta + d_2\tau, \\ \rho'\zeta' = a_3\xi + b_3\eta + c_3\zeta + d_3\tau, \\ \rho'\tau' = a_4\xi + b_4\eta + c_4\zeta + d_4\tau。\end{cases} \tag{1}$$

为了由此过渡到仿射群，我们先指出，若投影的结果使无穷大平面转换成其本身，即若每一个 τ 等于 0 的点对应于一个 τ' 等于 0 的点，则此投影变换为仿射变换。若 $a_4 = b_4 = c_4 = 0$，就出现这种情况。因此如果用 $\rho'\tau'$ 来除方程组(1)的每一个方程，以求出非齐次方程，并仍以 $a_1\cdots$ 来代替 $a_1 : d_4\cdots$，那么得

$$\begin{cases}x' = a_1x + b_1y + c_1z + d_1, \\ y' = a_2x + b_2y + c_2z + d_2, \\ z' = a_3x + b_3y + c_3z + d_3。\end{cases} \tag{2}$$

这实际上就是原有的仿射公式。由此可知，根据无穷远平面仍保持不变的条件，可从投影群 G_{15} 中分出 12 个参数的子群，即仿射群。

同样地，要求出主群 G_7，可以选出那些不仅可使无穷大平面，而且可使虚球面圆不变的投影变换(或仿射变换)，即在这种投影变换下，每一个满足方程 $\xi^2 + \eta^2 + \zeta^2 = 0$ 和 $\tau = 0$ 的点，对应于满足同样方程的点。这个论断是很容易证实的。你们只要记住，根据我们的

条件，可能差一个常数因子，借助于 $\tau' = 0$ 平面上的仿射变换而固定了对应于虚球面圆的二次曲线的 6 个（齐次）常数，因而在仿射变换的 12 个常数上有 $6 - 1 = 5$ 个条件，$12 - 5 = 7$，正好是 G_7 的参数。

1859 年，伟大的英国几何学家凯莱使这整个考虑方式发生了重要的改变[①]。在此以前，仿射几何和投影几何好像是度量几何的两个比较弱的部分，但凯莱相反地使我们有可能把仿射几何以及度量几何都看做是射影几何的特例。他说："投影几何就是全部几何。"这种从表面上来看似是而非的联系，是因为人们把所研究的图形与某种流形（即无穷大平面或其上的虚球面圆）结合在一起。因而一个图像的仿射性质或度量性质，不外乎是由此延伸的图形的射影性质。

让我们用两个非常简单的例子来说明这一点，以有所不同的形式提出众所周知的事实。两条直线平行的说法，在投影几何中没有直接的意义。但是如果我们对给定的流形（两条线）加上无穷远的平面，那么说两条给定的线在给定的平面上相交，这种说法就是纯粹的射影几何的说法。一条线垂直于一平面，情况也类似。我们可以把它化为给定图形的极线关系（投影性质），方法是加上虚球面圆使给定图形得到延伸（图 20.1）：线上点的图形 P_∞ 和平面上的线的图形 g_∞，在无穷大的平面上，对于虚球面圆是极点和极线。

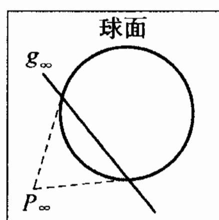


图 20.1

我想把这里已简单提示的思路发挥得更充分一些，说明怎样循此思路建立起一个完整的几何结构系统。在这一方面作出最大贡献的是英国数学家，我已经提到过凯莱，其次是都柏林大学的 J·J·西尔维斯特（J. J. Sylvester）和 G·萨蒙（G. Salmon）。他们从 1850 年开始创造了一门代数分支，狭义地称为线性齐次代换的

① 参见《关于四元数的第六篇专论》，伦敦皇家学会哲学学报，1859 年。也可看《数学文章集》第 2 卷，第 561 页及以后部分，剑桥，1889 年。

不变量理论[①]，有了这个理论，就有可能在凯莱原理的指导下，在分析的基础上，作出几何学的完整的系统结构。为了了解这个系统，我们有必要用一点时间来讨论不变量的理论。

§ 20.2 关于线性代换的不变量理论

当然我只能介绍主要的结果和思路，不谈细节和证明。至于这个广大领域的文献，我首先请你们参考 *Jahresberichte der deutschen Mathematiker-Vereingung*（1892 年）第一卷中 W·弗朗茨·梅耶写的"Die Fortschritte der projectiven Invariantentheorie in letzten Vierteljahrhundert"一文，以及同一作者在《百科全书》（IB2）中写的"不变量"一文。关于不变量理论特别涉及几何学的方面，在 G·萨蒙的教科书[②]里可以找到所需的一切材料，那本书对于传播这里所涉及的思想做出了最大的贡献。W·费德勒（W. Fiedler）所译的该书的德文本，始终得到特别广泛的采用。林德曼（F. Lindemann）编辑的 A·克莱布什的讲义[③]属于同一范畴的书。

（1）现在回到我们的正题。设有任一给定变量的数目，相应地有一个二元、三元、四元、……，区域，为了使我们能把前三者情况下的变量最终看成线、面或空间上的齐次坐标，我们用记号

$$\xi,\ \tau;\ \zeta,\ \eta,\ \tau;\ \xi,\ \eta,\ \zeta,\ \tau;$$

其中 $\tau=0$ 始终表示无穷远的元素。

---

① "不变量理论"这个术语也用于广义，指任意的变换群。我们在以后几页中用的是狭义，是西尔维斯特首先这样用的。

② G·萨蒙：《分析几何》，内容有：Ⅰ. 圆锥曲线；Ⅱ. 高次平面曲线；Ⅲ. 空间；Ⅳ. 线性变换代数讲义。德文译者 W·费德勒，莱比锡（杜伯纳出版社）。每一卷出了几版，卷Ⅰ新近由 F·丁格尔迪（F. Dingeldey）编辑；卷Ⅲ由 K·科默莱尔（K. Kommerell）及 A·布里尔（A. Brill）编辑。

③ A·克莱布什：*Vorlesungen über Geometrie*，F·林德曼编辑，莱比锡（杜伯纳出版社），1876 年第一版，1906 年第二版。

(2) 我们来考虑这些变量的一切齐次线性变换群。目前我们心中考虑的不仅有变量的比，正像以后在射影几何中的情况一样，而且有变量的个别值。我们可记这些代换为

$$\begin{aligned}\xi' &= a_1\xi + d_1\tau,\\ \tau' &= a_4\xi + d_4\tau,\end{aligned}\qquad \begin{aligned}\xi' &= a_1\xi + b_1\eta + d_1\tau,\\ \eta' &= a_2\xi + b_2\eta + d_2\tau,\\ \tau' &= a_4\xi + b_4\eta + d_4\tau,\end{aligned}\qquad \begin{aligned}\xi' &= a_1\xi + b_1\eta + c_1\zeta + d_1\tau,\\ \eta' &= a_2\xi + b_2\eta + c_2\zeta + d_2\tau,\\ \zeta' &= a_3\xi + b_3\eta + c_3\zeta + d_3\tau,\\ \tau' &= a_4\xi + b_4\eta + c_4\zeta + d_4\tau。\end{aligned}$$

在这 3 个群中的参数个数相应地为 4，9，16。

为了方便起见，我们按习惯在公式中只用变量 ξ 及 τ，只写出包含这两个变量的项，中间用点隔开。如果以后涉及二元区域，就去掉这些点号；对于三元或四元区域，我们用含 η 的项或含 η 和 ζ 的项来代替点号，这些项类似于已写出的项。因而，我们一般说变量是 ξ，…，τ，而线性代换是

$$\begin{cases}\xi' = a_1\xi + \cdots + d_1\tau,\\ \cdots\cdots\\ \tau' = a_4\xi + \cdots + d_4\tau。\end{cases}\tag{1}$$

(3) 至于不变量理论的对象，我们将以两种不同形式来考虑。第一种形式考虑的是变量的任一值组 ξ_1，…，τ_1；ξ_2，…，τ_2；ξ_3，…，τ_2；…，根据几何学的精神，可以直接记为点 1，2，3，…。这些值组的每一个将经受群(1)的代换，我们所关心的是建立在这些联立代换作用下仍然不变的值组的组合。

(4) 第二个形式除了考虑这样的点以外，还要考虑变量的函数，主要是有理整函数。事实上，我们可以只考虑齐次有理整函数(称为型)，因为，无论如何，由于代换的齐性，同次的项可以代换成原来一样的次数。因而我们将考虑线性型

$$\phi = \alpha\xi + \cdots + \delta\tau$$

及二次型

$$f = A\xi^2 + \cdots + 2G\xi\tau + \cdots + K\tau^2$$

等等。我们也可以同时考虑同次的几个型，这样就用下标来加以区别，即

$$\phi_1 = \alpha_1\xi + \cdots + \delta_1\tau;\ \phi_2 = \alpha_2\xi + \cdots + \delta_2\tau;\ \cdots。$$

同样地，我们可以从几个变量的型开始，即从双线性型开始：

$$f = A\xi_1\xi_2 + \cdots + G\xi_1\tau_2 + \cdots + N\tau_1\xi_2 + \cdots + P\tau_1\tau_2。$$

为了弄清楚这里出现的一般问题，我们必须首先问：在我们用群(1)来对这些变量施加变换，并规定型 ϕ 或 f 的值保持不变的情况下，这些型的系数如何变换？

先考虑线性型，令

$$\phi = \alpha\xi + \cdots + \delta\tau = \alpha'\xi' + \cdots + \delta'\tau'。$$

如果将(1)式表达的 ξ', …, τ'代进去，就得到变量 ξ, …, τ 的等式：

$$\begin{aligned}\alpha\xi + \cdots + \delta\tau &= \alpha'(a_1\xi + \cdots + d_1\tau) + \cdots + \delta'(a_4\xi + \cdots + d_4\tau)\\ &= (\alpha'a_1 + \cdots + \delta'a_4)\xi + \cdots + (\alpha'd_1 + \cdots + \delta'd_4)\tau。\end{aligned}$$

由此得

$$\begin{cases}\alpha = a_1\alpha' + \cdots + a_4\delta',\\ \cdots\cdots\\ \delta = d_1\alpha' + \cdots + d_4\delta',\end{cases} \tag{2}$$

因而线性型的新系数 α', …, δ' 用另一个线性代换同原有系数 α, …, δ 联系了起来。这另一个线性代换同(1)的关系是很简单的：系数矩阵的行和列相互交换（代换被“转置”过来），而且原有系数（不带撇）和新系数（带撇）相互交换其位置。这个新的代换(2)被称为原

代换(1)的逆步代换,简单地称作线性型系数 $\alpha, \cdots, \delta$ 逆步于变量 $\xi, \cdots, \tau$。受到同样代换(1)的变量集合 $\xi_1, \cdots, \tau_1; \xi_2, \cdots, \tau_2; \cdots$用类似的术语称为逆步变量。

现在来谈二次型 f。我们首先问各个二次项 $\xi^2, \cdots, \xi\tau, \cdots, \tau^2$,在线性代换(1)作用下的情况。从(1)式立刻对新变量的各二次项得出公式

$$\begin{cases}\xi'^2 = a_1\xi^2 + \cdots + 2a_1d_1\xi\tau + \cdots + d_1^2\tau^2, \\ \cdots\cdots \\ \xi'\tau' = a_1a_4\xi^2 + \cdots + (a_1d_4 + a_4d_1)\xi\tau + \cdots + d_1d_4\tau^2, \\ \cdots\cdots \\ \tau'^2 = a_4^2\xi^2 + \cdots + 2a_4d_4\xi\tau + \cdots + d_4^2\tau^2。\end{cases} \tag{3}$$

我们可以把这些关系简单地表示如下。变量的各二次项与变量本身同时受到齐次线性代换。这种代换可以立即从(1)式导出。由于 f 是这些二次项的线性型,我们可以重复前述的推理,推出系数 $A, \cdots, 2G, \cdots, K$ 所受到的变换,是线性齐次的,逆步于项 $\xi^2, \cdots, \xi\tau, \cdots, \tau^2$的代换(3),即 $A, \cdots, 2G, \cdots, K$ 和 $A', \cdots, 2G', \cdots, K'$ 之间的方程,由(3)式得出,如同(2)式由(1)式得出一样。

(5) 我们现在可以来概括不变量理论的一般问题。给出任一点集 1, 2, …并给出某些线性型二次型或更高次型 $\phi_1, \phi_2, \cdots, f_1, f_2, \cdots$,因而不变量是指坐标 $\xi_1, \cdots, \tau_1; \xi_2, \cdots, \tau_2; \cdots$的函数和系数 $\alpha_1, \cdots, \delta_1; \alpha_2, \cdots, \delta_2; \cdots; A_1, \cdots, K_1; A_2, \cdots, K_2; \cdots$的函数,它们在变量的线性代换(1)及我们刚才确定的相关的系数系统的代换下保持不变。要研究的是一切可能的不变量的集合。

“协变”和“逆变”这两个词有时用于表示不变量中特殊的类。如果变量 $\xi_1, \cdots, \tau_1; \xi_2, \cdots, \tau_2; \cdots$本身出现在不变式中,我们就说它是协变式;如果线性型的系数 $\alpha_1, \cdots, \delta_1; \alpha_2, \cdots, \delta_2; \cdots$出现在不变式中,我们就说是逆变式。因此,“不变量”一词仅限于既不包含坐标

ξ_1，…，又不包含 α_1，…等系数，而仅由二次或更高次型系数构成的式子。这两种情况之所以需要强调并加以对比，是因为以变量 ξ，…，τ 的集为一方，以变量 α，…，δ 的集为另一方，表现出某种互逆性质：如果对其中之一进行某线性代换，那么另一方就发生逆步代换，而不管我们从哪个集开始。因而我们可以通过适当的排列，由某一类不变式导出另一类相似的不变式。至于几何解释，这里显然有一个对偶原理的表达式，因为如果我们把 ξ，…，τ 看做是点坐标，那么 α，…，δ 就成为齐次线坐标或面坐标。不过，区分 ξ，…，τ 或 α，…，δ 是否真正出现在所说的式子中，当然没有根本意义。以后我们用"不变量"一词，一般指的是广义的。

(6) 我们现在从另一方向更明确地定义不变式的概念，以便有可能建立起严格的理论。以后我们把不变式仅仅看做是坐标和系数的有理函数，而且它们对每一个点的坐标以及对其中出现的每一个型的系数是齐次的。我们可以把每一个这样的有理函数表示为两个整有理齐次函数的商，并用它们本身来进行研究。既然分子、分母的公因子不改变商的值，那么在至此为止所采用的意义上，分子和分母就不必是不变量，而是在每个线性代换下都得到的一个因子。

可以证明这个因式仅取决于代换的系数，且它一定是代换行列式的幂：

$$r=\begin{vmatrix} \alpha_1 & \cdots & d_1 \\ \cdots & \cdots & \cdots \\ \alpha_4 & \cdots & d_4 \end{vmatrix}。$$

这样最终就归结于考虑那样一些给定量集的有理齐次整函数，它们在我们已经建立的变量及系数的线性代换下是乘以代换行列式的幂 r^λ。这些函数我们称为相对不变量，因为它们所经历的变化总是非本质的，且在 $r=1$ 时的一切代换下完全保持不变。指数 λ 称为不变量的权。作为对比，我们把至此为止记作不变量的那种函数称为绝对不变量。因此，每一个绝对不变量是两个同权的相对不变量的商。

(7) 据上述,我们实际上对于不变量理论已获得了一个系统的观点。最简单的相对不变量是对给定的变量集为最低次的多项式,进而到高次多项式。如果 j_1, j_2 是任意两个相对不变量,那么它们的幂的每一个乘积 $j_1^{k_1}\cdot j_2^{k_2}$ 也是相对不变量。因为如果把因子 r^{λ_1} 代入 j_1,把 r^{λ_2} 代入 j_2,那么除了因子 $r^{k_1\lambda_1+k_2\lambda_2}$ 以外,$j_1^{k_1}\cdot j_2^{k_2}$ 都会重新变成其自身。如果我们现在作出各乘以常因数的这样一些项之和:

$$\sum_{(K_1,K_2,\cdots)} c_{K_1,K_2,\cdots} j_1^{k_1}\cdot j_2^{k_2}\cdots,$$

并且如果各个被加项在变换后都确实是乘以 r 的同样的幂,即它们都有同样的权("等权"),那么我们显然又得到一个较高次的相对不变量,因为各项的因子可以放在连加号之前。

不变量理论的中心问题自然是能不能用这样的方法得出一切不变量,在各个给定的情况下,什么是最低次不变量的完备系统?亦即可按上述方式作出系统中元素的有理分式及多项式,使其包含所有相对不变量。不管怎样,主要的定理是:对于每有限个给定量,始终存在一个有限的"完备不变量系统",即存在有限个不变量。所有其他不变量是这有限个不变量的有理分式或多项式。在不变量的系统理论方面所取得的这些确定结果,应归功于德国数学家 P·哥尔丹(P. Gordon)和 D·希尔伯特。希尔伯特在德国《数学年刊》第 36 卷中发表的专题研究[①],是特别值得指出的。

现在我要提出几个简单的例子,也就是以后在几何学中要用到的那些例子,以便弄清楚我们一直在讨论的抽象推导。这里当然只是讲个概要,不是详细证明。

(1) 首先假设我们在二元区域内仅有一定数量的点:

$$\xi_1,\ \tau_1;\ \xi_2,\ \tau_2;\ \xi_3,\ \tau_3;\ \cdots,$$

---

① "Über die Theorie der algebiaischen Formen",《数学年刊》第 36 卷,第 473 页及以后部分,1890 年。

这里我们得出一个有趣的定理：最简单的不变量可以从这些坐标构成的二阶行列式得到，而且这些行列式构成完备的不变量系统。

由 1 和 2 两个点可以建立一个二阶行列式

$$\Delta_{12} = \begin{vmatrix} \xi_1 & \tau_1 \\ \xi_2 & \tau_2 \end{vmatrix}。$$

这实际上是变量的整有理函数，而且对(ξ_1, τ_1)和(ξ_2, τ_2)都是齐次的。如果把行列式乘法法则运用于计算：

$$\Delta_{1'2} = \begin{vmatrix} \xi_1' & \tau_1' \\ \xi_2' & \tau_2' \end{vmatrix} = \begin{vmatrix} a_1\xi_1 + d_1\tau_1, & a_4\xi_1 + d_4\tau_1 \\ a_1\xi_2 + d_1\tau_2, & a_4\xi_2 + d_4\tau_2 \end{vmatrix}$$
$$= \begin{vmatrix} a_1 & d_1 \\ a_4 & d_4 \end{vmatrix} \cdot \begin{vmatrix} \xi_1 & \tau_1 \\ \xi_2 & \tau_2 \end{vmatrix} = r \cdot \Delta_{12},$$

我们能立刻看出这个行列式的不变性质。这里不变量的权是 1。

同样，n 个点 1, 2, …, n 共有 $\frac{n(n-1)}{2}$ 个权为 1 的不变量：

$$\Delta_{ik} = \begin{vmatrix} \xi_i & \tau_i \\ \xi_k & \tau_k \end{vmatrix} \quad (i, k = 1, 2, \cdots, n)。$$

要证明这些行列式构成完备的不变量系统，即 n 个点的每一个相对不变量可以表示为等权项的和：

$$\sum C \cdot \Delta_{ik}^{t} \Delta_{lm}^{t} \cdots,$$

那就走得太远了。我们从分子、分母为等权的相对不变量的商，求出最一般的有理绝对不变量。因此，最简单的绝对不变量的一个例子就是商$\frac{\Delta_{ik}}{\Delta_{lm}}$。

结合这个例子，我想解释一下在理论上非常重要的抽象概念，即“合冲”(即不变量的一个耦合或联系)。也就是说，某些基本不变量的组合可能等于零。例如，对于 4 个点，我们有

$$\Delta_{12}\Delta_{34}+\Delta_{13}\Delta_{42}+\Delta_{14}\Delta_{23}=0。$$

这不过是一个已知的行列式恒等式，实际上前面我们已经使用过它，在一个完备的系统的不变量之间的这种恒等式，就叫做合冲。如果有几个那样的合冲，就可以用相乘或相加的方法形成新的合冲，因而可以像对于行列式本身一样，可以问是否有能用这种方法形成其他一切"合冲"的完备的合冲系统问题。理论表明，总是存在那样一个有限的系统。例如，在 4 个点的情况下，这个完备的系统由上述一个方程构成，即在 6 个行列式 Δ_{12}，…，Δ_{34} 之间的一切恒等式都可由上述方程推出。在 5 个及 5 个以上点的情况下，完备系统由这种类型的一切方程构成。了解这些合冲，对于了解整个不变量系统当然有根本的意义，因为如果最简单的不变量的两个等权集合的差别仅为以合冲的左边作为因数的项，那么它们就是恒等的，不需要计算两次。

(2) 同样地，如果在三元或四元区域里有若干单个的点，那么整个不变量系统就是以完全一样的方式由它们的坐标形成的三阶行列式或四阶行列式构成。例如，在三元区域，3 个点的基本不变量还是权为 1 的不变量：

$$\Delta_{123}=\begin{vmatrix}\xi_1 & \eta_1 & \tau_1\\ \xi_2 & \eta_2 & \tau_2\\ \xi_3 & \eta_3 & \tau_3\end{vmatrix}。$$

其他一切细节，特别是这里怎样建立合冲，留给你们自己去考虑。

(3) 现在让我们来考虑四元区域中的二次型：

$$f=A\xi^2+2B\xi\eta+C\eta^2+2D\xi\zeta+2E\eta\zeta+F\zeta^2+2G\xi\tau$$
$$+2H\eta\tau+2J\zeta\tau+K\tau^2。$$

我们可以立刻写下一个仅依赖于 10 个系数 A，…，K 的不变量，即行列式

$$\Delta = \begin{vmatrix} A & B & D & G \\ B & C & E & H \\ D & E & F & J \\ G & H & J & K \end{vmatrix}。$$

因为系数 $A, \cdots, K$ 逆步地变换为 $\xi, \cdots, \tau$ 的二次项，所以很容易证明这个不变量的权是 -2，$\Delta' = r^{-2} \cdot \Delta$ 仅由这个型的系数形成的整个不变量系统，只包含这个 Δ，即每一个仅包含 $A, \cdots, K$ 的整有理不变量是 Δ 的幂的倍数。

如果我们现在把点的坐标 ξ, η, ζ, τ 加在前面的系数上，那么最简单的共同不变量或共变式(据上述术语)就是型 f 本身，因为系数 $A, \cdots, K$ 的变换完全是由系数不变性的规定所决定的。因此，每一个给定的型当然是它本身的共变式。事实上，根据定义，它在我们的代换作用下是完全不变的，因此是权为 0 的不变量或绝对不变量。此外，如果我们运用两个点 $\xi_1, \cdots, \tau_1$ 和 $\xi_2, \cdots, \tau_2$，就会出现作为新的共变式的所谓极型

$$A\xi_1\xi_2 + B(\xi_1\eta_2 + \xi_2\eta_1) + C\eta_1\eta_2 + \cdots + K\tau_1\tau_2。$$

其权又为 0，即它同样是绝对不变量。

最后，如果我们与 f 同时考虑线性型 ϕ，即其全体系数 $\alpha, \beta, \gamma, \delta$，即得以下权为 -2 的联立不变式，它是通过所谓用 $\alpha, \beta, \gamma, \delta$ 对行列式“加边”而产生的

$$\begin{vmatrix} A & B & D & G & \alpha \\ B & C & E & H & \beta \\ D & E & F & J & \gamma \\ G & H & J & K & \delta \\ \alpha & \beta & \gamma & \delta & 0 \end{vmatrix}。$$

根据前述，我们也可以称它为反变式。正如你们所知道的，这个行列式在解析几何中起着重要的作用。我们认识到，形成不变量的纯解

析过程，在这里是根本性的。

如果我们有包含系数 $\alpha_1, \cdots, \delta_1$ 和 $\alpha_2, \cdots, \delta_2$ 的两个线性型 ϕ_1，ϕ_2，则通过同一行列式的两重“加边”得另一个不变量

$$\begin{vmatrix} A & B & D & G & \alpha_1 & \alpha_2 \\ B & C & E & H & \beta_1 & \beta_2 \\ D & E & F & J & \gamma_1 & \gamma_2 \\ G & H & J & K & \delta_1 & \delta_2 \\ \alpha_1 & \beta_1 & \gamma_1 & \delta_1 & 0 & 0 \\ \alpha_2 & \beta_2 & \gamma_2 & \delta_2 & 0 & 0 \end{vmatrix}。$$

它同样有权－2。

以上这几点说明足以使你们大致了解涉及面甚广的不变量理论领域。这里发展的理论，其涉及面之广是非同一般的，而且运用了许许多多巧妙想法，特别是在为建立一个完备的不变量系统及对给定基本型建立一个完备的合冲系统的时候。下面让我再作一个一般特性的说明。在我们所举的例子中，我们总是通过建立行列式来得出不变量，这一点是我们把行列式理论作为不变量理论基础的根据。由于这种联系，凯莱起先对不变量用了超行列式的名称。引入“不变量”一词的是西尔维斯特。行列式的重要性，是很值得提出来的一个问题，因为在整个数学领域里，应该专门写一章的是行列式。凯莱有一次在谈话中对我说，如果要他就整个数学作 15 次讲演的话，他就要用一讲来讨论行列式。你们如果愿意的话，不妨根据自己的经验回想一下你们对行列式理论的价值是否有这样高的估计。在我自己讲的初等数学课中，我发觉自己为了教学法上的理由一直把行列式放在越来越突出的地位。我这样的经验实在太多了，学生虽然掌握了对于简化长长的表示式非常有用的公式，却常常不能弄清公式的意义，他们只顾计算技巧的熟练，不深入探究所学的内容，因而也妨碍他们达到真正的掌握。从总的方面来考虑，也从所研究的不变量

理论来说，行列式当然是必不可少的。

下面我们转入正题，借助于上述考虑，使几何学系统化。

§ 20.3　不变量理论在几何学上的应用

我们先用变量 ξ, …, τ 来表示一般非齐次直角坐标：(ξ, τ)为平面坐标，(ξ, η, τ)为三维空间坐标，(ξ, η, ζ, τ)为四维空间坐标，等等。因而不变量理论的线性齐次代换

$$\begin{aligned} \xi' &= a_1\xi + \cdots + d_1\tau, \\ &\cdots\cdots \\ \tau' &= a_4\xi + \cdots + d_4\tau \end{aligned} \tag{1}$$
[①]

表示带有固定坐标原点的空间仿射变换的全体。除了可能加一个因子外，每一个相对不变量本身将是一个不因这些仿射变换而变化的几何量，即在这些变换所定义的仿射几何中有确定意义的量。

例如，如果在二维情况下（即在平面上）给定了两个点 1 和 2，则如我们所知，基本不变量 Δ_{12} 表示有适当符号的三角形(012)面积的两倍。事实上已经知道，一个仿射变换只是把三角形面积乘以代换的行列式，这正意味着 Δ_{12} 是权为 1 的一个相对不变量。两个面积的商 $\frac{\Delta_{12}}{\Delta_{34}}$ 保持绝对不变，但方程 $\Delta_{12}=0$ 也如此，因为在这个方程中乘以因子是没有关系的。实际上，这个方程对于我们所说的仿射变换具有绝对不变的意义，即 3 个点 0, 1, 2 处在同一直线上。

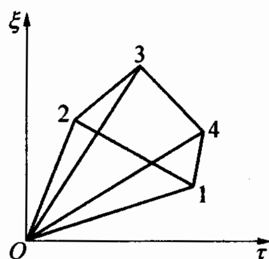


图 20.2

如果我们有几个点 1, 2, 3, …（图 20.2），那么完备不变量系统就由这些点

① 式码为编辑所加。

的一切行列式 Δ_{ik} 所构成。因而如果能构造一个量，它是坐标的有理整函数，并在一切仿射变换(1)下是相对不变的，即在我们讨论的仿射几何中总是有它的意义，那么它一定能表示为 Δ_{ik} 的多项式。对于简单的情况，例如，平面上的每一面积[如多角形(1, 2, 3, 4)的面积]，可以立刻用几何方法证实是这种不变量，而我们在前面给出的多边形面积一般公式

$$(1,\ 2,\ 3,\ 4) = \Delta_{12} + \Delta_{23} + \Delta_{34} + \Delta_{41},$$

实际上就只不过是一般定理对于这个特例的表示式。

最后让我们来考虑不变量之间的合冲。基本的合冲

$$\Delta_{12}\Delta_{34} + \Delta_{13}\Delta_{42} + \Delta_{14}\Delta_{23} = 0$$

表示着由 4 个任意点和原点构成的 6 个三角形面积之间的恒等式，因而是我们所说的仿射几何的一般定理。对于每一个合冲，当然有类似结果。反过来说，我们的仿射几何的每一个定理，只要是仿射变换(1)的不变量之间的关系，就必须以一个合冲来表示。因此，根据前面关于 4 点情况下有一完整合冲系统的论断，可以说对于 4 点系统成立的仿射几何的一切定理一定可由刚才给出的那个一般定理导出。我们可以用同样的方法确定下述一般论断的正确性，由于不变量理论提供了不变量和合冲的完备系统，因此由它可毫无例外地系统列举所有可能的量和定理。

这里我要略去这个问题的细节不谈了，我只想提一下，和点相联系，还可以考虑由型 $\phi = \alpha\xi + \delta\tau$, $f = A\xi + 2G\xi\tau + K\tau^2$, … 所决定的几何流形。这样的型使平面上的每个点和一个数之间建立了对应关系，即决定着一个标量场。带着这个观点，我们就很容易从几何上解释一个给定型的不变量，而且不变量之间的每个合冲又将表示一个几何定理。

至此为止，我们所考虑的，可称之为 n 维几何中的不变量理论的直观解释，其中 n 个变量被认为是直角坐标。除了这个解释以外，还

有一个本质上不同的解释：可以把变量看做为 $(n-1)$ 维空间 R_{n-1} 中的齐次坐标，它的非齐次坐标是 $x=\frac{\xi}{\tau},\cdots$，其中 n 个坐标的公因子是非本质的。我们在前面讨论过 R_{n-1} 和 R_n 中坐标之间的联系。我们曾把 R_{n-1} 看做为 R_n 中的线性 $(n-1)$ 维流形 $\tau=1$，并通过 R_n 的原点的射线来投射其各个点。因此，R_{n-1} 中一个点的所有可能的齐次坐标值组的集合，等于对应于它的 R_n 中点的坐标值的集合。现在，R_{n-1} 中齐次变量的线性代换表示投影变换。事实上，形为

$$\begin{aligned}&\rho'\xi'=a_1\xi+\cdots+d_1\tau\\&\cdots\cdots\\&\rho'\tau'=a_4\xi+\cdots+d_4\tau\end{aligned}$$

彼此之间只差一个任意因子 ρ' 的所有代换，产生同样一个射影变换。所有这些射影变换的群所包含的，不是 n^2 个任意常数，而仅仅是 n^2-1 个任意常数；特别是在 R_2 和 R_3 中，那样常数的数目分别为 8 和 15。

不过如果我们想按 R_{n-1} 的投影几何解释 n 个变量 $\xi,\cdots,\tau$ 的不变量理论，那么必须首先记住，正因为我们用的是齐次坐标，所以在不变量理论中，可能解释的只是这样的数量和关系：其中出现的每一个点的坐标 $\xi,\cdots,\tau$ 是零阶齐次的，并且其中出现的每个线性型、二次型或其他型的系数组也具有同样的性质。

用具体的例子来讲就清楚了。只要讨论二元区域 $(n=2)$ 就行了。假设有两个变量 ξ 和 τ，把 $x=\frac{\xi}{\tau}$ 看做是直线上的横坐标。如果给出一系列值组 $(\xi_1,\tau_1),(\xi_2,\tau_2),\cdots$ 则行列式

$$\Delta_{ik}=\begin{vmatrix}\xi_i&\tau_i\\\xi_k&\tau_k\end{vmatrix}\qquad(i,k=1,\cdots,p)$$

表示整个基本不变量系统。在所有关于不变量的陈述中，哪一些具

有射影几何上的意义呢？关于 Δ_{ik} 之一具有某一确定数值的陈述当然没有投影几何上的意义，因为如果把 ξ_i，τ_i 乘上因子 ρ（这不会改变点 i），那么我们也把 Δ_{ik} 乘上 ρ。但是，Δ_{ik} 等于 0，即关系 $\Delta_{ik}=0$，却有射影几何上的意义，因为我们可以写它为 $\frac{\xi_i}{\tau_i}=\frac{\xi_k}{\tau_k}$，从而实际上只出现点坐标之比，其几何意义是点 i 和点 k 重合，这是显而易见的。

现在为了求得对每个点的坐标本身是零次的数值不变量，我们必须把两个以上的点组合起来。试验表明我们至少需要 1，2，3，4 共 4 个点。此时，型

$$\frac{\Delta_{12}\cdot\Delta_{34}}{\Delta_{14}\cdot\Delta_{32}}$$

的每个商对 4 对变量 (ξ_1,τ_1)，(ξ_2,τ_2)，…，(ξ_4,τ_4) 的每一对是零维齐次的。由此可见它的权是 0，也就是说，它是一个绝对不变量。于是，这个量具有投影几何的意义，并表示着对直线上一切投影变换都是不变的一个数值。它当然只不过是有确定顺序的 4 个点的交比。因为按非齐次坐标可以记为下列形状

$$\frac{x_1-x_2}{x_1-x_4}:\frac{x_3-x_2}{x_3-x_4}。$$

从不变量理论的观点，线上点列的最简单的不变量，就是我们求得的 4 个点的交比，它满足为使不变量得到投影几何意义所必要的齐次性条件。

这里还想谈一点一般的意见。多年来我一直在思考投影几何中的一个普遍倾向，即从交比出发来解一切表现出不变性质的量。但从我们已经取得的结果来看，可以得出这样一个判断：从这个方向去努力，只会使我们更难以深入理解射影几何的结构。最好是先找一切有理整式（相对）不变量，由这些不变量首先构成有理不变量，特别是绝对有理不变量，从中再构成满足投影几何齐次性条件的不变

量。这种方式,就是由简单到复杂的循序渐进的方式。如果我们把一个特殊的有理不变量——交比放在首位,企图完全由这个不变量去构成其他不变量,那么这个程序就混淆不清了。

现在我们来看从不变量 Δ_{ik} 之间的合冲能推导出哪一类射影几何的定理。我们从下列基本合冲出发

$$\Delta_{12}\Delta_{34}+\Delta_{13}\Delta_{42}+\Delta_{14}\Delta_{23}=0,$$

用左侧的最后一个被加数除全式,注意到 $\Delta_{23}=-\Delta_{32}$, $\Delta_{24}=-\Delta_{42}$,得

$$\frac{\Delta_{12}\Delta_{34}}{\Delta_{14}\Delta_{32}}=1-\frac{\Delta_{13}\Delta_{24}}{\Delta_{14}\Delta_{23}}。$$

这样,根据原来的定义我们在左侧得出点 1, 2, 3, 4 的交比。在右侧,我们得出在 2 和 3 的顺序改变后以同样方式构成的同样点的交比。如果我们除以其他的项,那么就得出其他顺序的交比。因此,4 个点的不变量之间的基本合冲之所以获得它们的几何意义,在于根据 4 个顺序而使交比所取得的 6 个值之间的已知关系。

我在这里不再进一步细说直线的投影几何是怎样建立在这个基础上的,同样也不再细说平面和空间投影几何中三元及四元不变量理论的解释是怎样推导出来的。详细资料可以在已经提到过的萨蒙—费德勒和克莱布什—林德曼的书中找到,那两本书中连续使用的正是对于不变量理论所作出的这种解释,因而射影几何得到了独立完整的表述,不仅是对于可以考虑的量(对应于不变量),而且对于可以建立的定理(对应于合冲),都是如此。不过说实在的,这个解释对于几何学家是可以了,对于研究不变量的人却不够。对于他们,在 R_{n+1} 的仿射几何研究中给出的解释更有价值。因为在 R_n 中,正如我们已经知道的,只有满足齐次性条件的那些不变量和合冲才是有用的。

下一节我想比较详细地讨论特别重要的一点,以便恢复前面被

打断的讨论。我想要说明凯莱原理是怎样在投影几何内使用不变量理论对仿射几何和度量几何进行分类的。

§20.4 凯莱原理和仿射几何及度量几何的系统化

我们在这里所讨论的是一般仿射几何，不像开始讨论关于不变量理论的完整解释时那样，假定有一个特殊的固定点——坐标的原点。

对于三维空间，我们立刻从非齐次坐标 x, y, z 开始，或在可能的情况下，从齐次坐标 ξ, η, ζ, τ 开始。于是，根据凯莱原理，当我们对给定的流形加上无穷远平面 $\tau=0$，或除这个平面外还加上虚球面圆 $\tau=0$, $\xi^2+\eta^2+\tau^2=0$，则可以从射影几何分别导出仿射几何或度量几何。

关于虚球面圆的一个说明可简化以下的讨论。我们在这里已经用了两个方程定义虚球面圆，即定义它为无穷远平面和通过原点的一锥面的相交线。如果我们把它看做是与之相切的一切平面的包络，那么就可以用一个平面坐标方程来确定它，或实际上确定了任意一锥面。如果像前面一样用 α, β, γ, δ 来记“平面坐标”，即线性型 ϕ 的系数，那么很容易证明，虚球面圆的方程是 $\alpha^2+\beta^2+\gamma^2=0$。换句话说，这个方程是使平面 $\alpha\xi+\cdots+\delta\tau=0$ 切于虚球面圆的条件。

现在就不难理解怎样用不变量理论相应地过渡到仿射几何及度量几何了。对于给定值组——描述所讨论的图形的点坐标、线性型及二次型等，我们分别加上确定的线性型 τ(即系数组 0, 0, 0, 1)，或用平面坐标写出的二次型 $\alpha^2+\beta^2+\gamma^2$。如果我们完全像前面一样处理这样扩展的型的系统，即如果我们建立全部不变量及不变量之间的合冲的系统，并强调其中满足齐次性条件的系统，那么对于最初给出的元素就分别得出了仿射几何及度量几何的一切概念及定理。

这样，不变量理论的研究结果，就转移到仿射几何及度量几何上。我想再次提请你们注意这样一个事实，对有理整式不变量及合冲特别加以强调，就可以对几何形成一个系统的观点，否则几何系统是搞不清楚的。

下面我不发抽象议论了，我想通过实例来说明我们怎样把仿射几何和度量几何最初步的基本量分别同时表示为给定量的系统及型 τ 和 $\alpha^2+\beta^2+\gamma^2$ 系统的不变量，以便立即把这些关系搞清楚。

首先我从仿射几何中取由 4 点构成的四面形的体积 T 作为例子。正如你们所知道的，这可以由下列公式表示：

$$T=\frac{1}{6}\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{vmatrix}=\frac{1}{6\tau_1\tau_2\tau_3\tau_4}\begin{vmatrix} \xi_1 & \eta_1 & \zeta_1 & \tau_1 \\ \xi_2 & \eta_2 & \zeta_2 & \tau_2 \\ \xi_3 & \eta_3 & \zeta_3 & \tau_3 \\ \xi_4 & \eta_4 & \zeta_4 & \tau_4 \end{vmatrix}。$$

现在我们要问：这个式子在何种程度上肯定了不变量性质？首先我们知道，这个行列式实际上是 4 个点的基本相对不变量（见前一节）。其次，在这 4 个点的分母中，有伴随我们图形的线性型 τ 的值，这些值是可以用一个型（见前一节）来作出的非常简单的（绝对的）不变量。这当然是意味着，在变换以后，线性型 τ 所变换成的型的值要记在分母中。或者说，如果一般地与型 $\alpha\xi+\beta\eta+\gamma\zeta+\delta\tau$ 相联系，那么这个型在点 1，…，4 的 4 个值的乘积要出现在分母中。因而 τ 本身也是一个有理不变量，而且对 4 个点的每个点的坐标，它确实是零维齐次的。说得确切一些，对于出现在分母中的伴随线性型 0，0，0，1（或可以说是 α，β，γ，δ）的系数，T 的维是 -4。既然这些量的公因子是任意的，则 T 的绝对值在射影几何中不可能有意义。事实上，除非像我们使用非齐次坐标时始终做的那样，已经选一单位线段或单位四面体，否则在仿射几何中没有办法把一定的数值赋予四面体的体积。但从我们现在总的观点来说，这就意味着，除了“无限远的

平面”$\tau=0$ 以外,我们的图形上应加上其他元素。举例来说,如果我们加上第五个点,取类似于 T 的两个表达式的商,那么我们实际上就得出了一个满足一切齐次性条件的表达式。因而这个式子必为仿射几何的一个绝对不变量,单一的式子 T 是仅有的权为 1 的相对不变量,正如前已确知的那样(见“分析的定义及基本性质”一节)。

这里我们又要回顾本卷的第一部分,因为那一主要部分的意义现在才显得比较明确。我们在仿射几何的专门研究部分中已认识到,我们在那一部分所推导出来的格拉斯曼基本几何量,完全属于仿射几何的范围。但是提供这些几何量的格拉斯曼行列式原理,绝不是杂乱无章的手段。相反,正如我们现在可以看到的,它是不变量理论在仿射几何(即射影几何与无穷远平面的结合)中完全自然的应用。之所以出现一般行列式——线段、面积、体积,可以由刚才讨论过的例子得到充分的解释。还有待说明的是:怎样由不变量理论导致由矩阵的子式所定义的一般格拉斯曼元素。这又要用例子来说明。给出平面上的两点(ξ_1, η_1, τ_1)及(ξ_2, η_2, τ_2),我们希望在仿射几何流形(线段、直线……)的不变量理论中求出这两点的等价物。这与前已获得的结果是严整地一致的,如果加上第三个“未确定”点(ξ, η, τ),并把基本不变量

$$\frac{1}{\tau\tau_1\tau_2}\begin{vmatrix} \xi & \eta & \tau \\ \xi_1 & \eta_1 & \tau_1 \\ \xi_2 & \eta_2 & \tau_2 \end{vmatrix}$$

看做是 ξ, η, τ 的线性型的话。这些变量的 3 个系数,即矩阵

$$\frac{1}{\tau_1\tau_2}\begin{pmatrix} \xi_1 & \eta_1 & \tau_1 \\ \xi_2 & \eta_2 & \tau_2 \end{pmatrix} \text{或} \begin{pmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{pmatrix}$$

的行列式因而成为关于新定义的流形的特征的量,实际上正好导致前已用以定义线段 1, 2 的矩阵。对于空间情况,我们可以用完全相同的方式,分别加上一个或两个未确定的 4 个坐标所成的组,由 3 点

或两点建立一个相对不变量的线性型或双线性型，它们的系数又提供一个平面片或一个空间线段的坐标，这与旧的定义完全一致。除了这些看法以外，我不再详谈了，这些看法大约已足以成为进一步研究的基础。

格拉斯曼原理在不变量理论中的地位既已说明，现在必须提出关于它的用途问题。在这一方面，应特别同前述的分类原理加以比较。那些分类原理是针对主群的特殊情况，并向我们列出了一切基本的几何流形。很明显，可以适当地把分类原理引申到任意线性变换群的情况。根据此原理，在每一种“几何”中，一方面固然要考虑至今为此提供不变量的给定系列的量(坐标、型系数等)的单个有理整函数，同时也要考虑那样的函数组 Ξ_1, Ξ_2, …。如果那样一个函数组在有关的群的一切代换下变换成它本身，即如果变换后的系列量按类似方式构成的函数 Ξ_1', Ξ_2'，能借助于一个确定的和唯一的由基本转换函数得出的系数，并只用 Ξ_1, Ξ_2, …线性地表示出来，那么我们说那个函数组定义了所指的几何的流形。组成函数组的每个函数，称为流形的分量。决定一个几何流形性质的特性，是流形的分量在所考虑的群的变换下的特点。当两个几何流形的分量构成两组同样数量的表示式，其中每一个表示式在坐标改变下发生同样的线性代换，即按我们前面用的术语它们是逆步的时候，我们说这两个几何流形是同类的。如果定义一个几何流形的函数组由单个函数构成的，那么线性代换就归结为乘以因数，这个函数就是一个相对不变量。

这些讨论比较抽象，我就从三元域的不变量理论中取一个简单例子来把它讲明白。关于三元域的不变量理论，我将在带固定原点的三维空间仿射几何中加以解释。如果给定两点(ξ_1, η_1, τ_1)及(ξ_2, η_2, τ_2)，那么两个三维坐标齐次对称出现的最简单的函数组，是 9 个双线性项的函数组：

$$\xi_1\xi_2,\ \xi_1\eta_2,\ \xi_1\tau_2,\ \eta_1\xi_2,\ \cdots,\ \tau_1\tau_2。\tag{1}$$

在线性变换下，用我们的习惯记号得

$$\begin{cases}\xi_1'\xi_2' = a_1^2\xi_1\xi_2 + a_1b_1(\xi_1\eta_2 + \eta_1\xi_2) + \cdots + d_1^2\tau_1\tau_2,\\ \xi_1'\eta_2' = a_1a_2\xi_1\xi_2 + a_1b_2\xi_1\eta_2 + a_2b_1\eta_1\xi_2 + \cdots + d_1d_2\tau_1\tau_2,\\ \cdots\cdots\\ \tau_1'\tau_2' = a_4^2\xi_1\xi_2 + a_4b_4(\xi_1\eta_2 + \eta_1\xi_2) + \cdots + d_4^2\tau_1\tau_2。\end{cases} \tag{2}$$

也就是说，这 9 个量的型正是我们刚刚讨论过的那种函数组，我们将把它们看做是仿射几何中一个流形的决定元素。这样的流形，以及任何根据方程(2)变换的 9 个量组成的其他函数组，称为张量。

观察方程(2)时，我们注意到，我们可以从量(1)中的 9 个量中一方面导出 6 个简单的线性组合；另一方面导出 3 个简单的线性组合，这些组合在线性代换下变换成它的本身。事实上，如果我们把量(1)排列成二次组

$$\begin{matrix}\xi_1\xi_2 & \xi_1\eta_2 & \xi_1\tau_2\\ \eta_1\xi_2 & \eta_1\eta_2 & \eta_1\tau_2\\ \tau_1\xi_2 & \tau_1\eta_2 & \tau_1\tau_2\end{matrix}$$

第一个集是对称于对角线的各项之和

$$2\xi_1\xi_2,\ \xi_1\eta_2 + \eta_1\xi_2,\ \xi_1\tau_2 + \tau_1\xi_2,\ \cdots,\ 2\tau_1\tau, \tag{3}$$

另一个集是它们的差

$$\xi_1\eta_2 - \eta_1\xi_2,\ \xi_1\tau_2 - \tau_1\xi_2,\ \eta_1\tau_2 - \tau_1\eta_2, \tag{4}$$

立刻可以从方程(2)得出函数组(3)和函数组(4)的代换公式。这样，我们已为仿射几何得出了两个新的流形，其中之一由 6 个量构成流形(3)，称为对称张量，而由 3 个量构成的流形(4)是我们已知的平面片。这个名称当然可以用于任一同步变换量的系统。下面我们直接说明使用“对称的”这个形容词的理由。

至于 3 个变量(4)的几何意义，我们知道，3 个量是由坐标点

(ξ_1, η_1, τ_1), (ξ_2, η_2, τ_2)和原点构成的三角形在坐标平面上的射影的两倍,每一个三角形周线按恰当的方向走过。我们这里所得到的,正是格拉斯曼行列式原理导出的最初几个流形之一。因此,我们可以阐明下述定理:通过我们的分类原理对仿射几何流形进行系统探索,必然使我们得出格拉斯曼行列式原理以及由此原理决定的几何流形。当然,这一点我在这里不再细谈。我只要说明,如果我们通过凯莱原理、通过四元不变量理论,用类似的方法对待一般仿射几何,那么就可以导出一切流形。

但是我们这个研究所获得的最重要结果,是了解到格拉斯曼行列式原理只是某种特殊的东西,它本身并不产生仿射几何的一切流形。我们用张量(1)和张量(3)所求得的流形,就是完全新的几何流形。

由于这些流形对于物理学的许多领域具有重大的意义,如弹性变形及相对论,因此我要简短地加以讨论。首先我要就这些量的名称讲点意见,这些意见应有助于读者弄清关于张量计算的新文献。

我在本书第一卷中讨论哈密顿四元数计算的时候用了“张量”这个词,但意义与现在所用的张量不同。如果 $q = a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k}$ 是一个四元数,我们称表示式 $T = \sqrt{a^2 + b^2 + c^2 + d^2}$ 是四元数的张量。哈密顿引进这个名称是有道理的,因为可以把乘以四元数用几何方式解释为从固定的原点旋转和伸展,这一点我们在第一部分已作了充分的解释。伸展的度量原来正好是我们称为张量的平方根 T。W·伏格特(W. Voigt)在他的晶体物理著作①中所用的张量一词,就同这个用法有密切关系。伏格特用张量这个词来记对应于事件的有定向的量,如杆的纵向伸展和压缩,杆端的推拉力施加于杆轴

① 例如,请看:(a)“Der gegenwärtige Stand unserer Kenninisse der Kristallelastizität”;(b)“Über die Parameter der Kristallphysik und über gerichtete Grössen höherer Ordnung”。两篇专论都刊登在 *Göttinger Nachrichten*,1900 年。

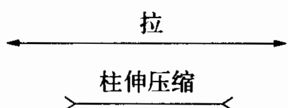
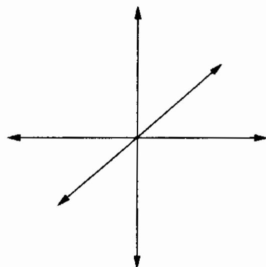


图 20.3

方向，但方向相反。我们可以用一个线段图来表示这样一个张量，线段顶端带有两个方向相反的箭头(图 20.3)。

我们可以把这样指定的一个张量的方向特征记为"双向的"，而与之相对比，把矢量记作"单向的"。这种张量常作为三重张量出现在物理中，即 3 个张量互成直角(图 20.4)。前面我们曾提到纯应变(纯仿射变换)，把它当作空间在具有固定原点的 3 个互相正交方向的均匀伸展。现在我们可以说，一个纯应变用一个三重张量来做几何表示，以代替以前的说法。如果我们把空间的那种沿 3 个方向伸展的概念看做是单个的几何量，略去"三重"一词，称这个量是张量，那么这个"张量"就是通常意义上的张量。这个意义上的张量的概念，正是我们前面所称的"对称张量"。


图 20.4

实际上，具有固定原点的纯应变由下述代换给出：

$$\begin{cases}\xi = a_{11}x + a_{12}y + a_{13}z, \\ \eta = a_{12}x + a_{22}y + a_{23}z \qquad (a_{ik} = a_{ki}), \\ \tau = a_{13}x + a_{23}y + a_{33}z,\end{cases} \tag{5}$$

我们把 3 个数的组(x, y, z)和(ξ, η, τ)解释为同一个直角坐标系中的点的坐标。变换的系数矩阵对于主对角线是对称的。如果我们现在转到具有同样原点的一个新的直角坐标系，那么正如简单计算所表明的，我们对所说的应变得出下列新的表示式

$$\begin{cases}\xi' = a'_{11}x' + a'_{12}y' + a'_{13}z', \\ \eta' = a'_{12}x' + a'_{22}y' + a'_{23}z' \qquad (a'_{ik} = a'_{ki}), \\ \tau' = a'_{13}x' + a'_{23}y' + a'_{33}z'。\end{cases} \tag{6}$$

根据同样公式得出 x, y, z 和 x', y', z'之间的关系就是 ξ, η, τ 和 ξ', η', τ'之间的关系。对于 a'_{11}, a'_{12}, …, a'_{33} 这 6 个系数,我们发现

(1) 它们线性地依赖于且仅依赖于这 6 个系数 a_{11}, a_{12}, …, a_{33},即它们定义一几何流形。

(2) 它们和我们在前面指定作为对称张量的分量的,关于坐标是双线性的表达式(3)的变换完全一样地变换。用"对称"这个形容词的根据是变换公式(5)和(6)中的系数矩阵的形式。

现在让我们转到一般仿射变换上

$$\begin{cases}\xi = a_{11}x + a_{12}y + a_{13}z,\\ \eta = a_{21}x + a_{22}y + a_{23}z,\\ \tau = a_{31}x + a_{32}y + a_{33}z,\end{cases} \tag{7}$$

其中原点是固定的。然后,按刚才讲的相应方法展开,即在正交变换几何中,9 个系数 a_{11}, a_{12}, a_{13}, …, a_{33} 完全与坐标乘积(1)一样变换。因此它们构成同类的量的分量。根据我们的术语,"张量"一词并不专门限于纯应变,所以上面那段话等于是说,一般仿射变换的系数矩阵是一个张量。

数学文献中存在着大量表示这个概念的其他名称。某些最通用的名称如下:

(1) 仿射向量(因同仿射变换有关)。

(2) 线性矢量函数(因为线性代换(7)可以这样解释:对于从原点开始的矢量 x, y, z,另一个类似的矢量 ξ, η, τ 可以通过代换(7)与之相对应)。

(3) 并向量和双积。但是第一个术语最初用于特殊的情况,以后再解释。

平面片(4)的分量也可以看做为一个变换的系数,即

$$\begin{cases}\xi = 1\cdot x - c\cdot y + b\cdot z,\\ \eta = c\cdot x + 1\cdot y - a\cdot z,\\ \tau = -b\cdot x + a\cdot y + 1\cdot z\end{cases} \tag{8}$$

类型的变换。确实不难说明,这个代换的系数在直角坐标变换下的性质和双线性表示(4)式一样。由于变换(8)中的系数矩阵的结构如此(对于主对角线,在改变符号后是对称的),因此由它决定的量也称为反对称张量。

从几何上来说,(7)式可以解释为通常的均匀形变,(6)式可以解释为纯形变(无旋转),(8)式可以解释为无穷小旋转。因此一个均匀无穷小形变之分解为一个纯形变及一个旋转,在直观上对应于我们从坐标乘积(1)导出对称张量(3)及反对称张量(4)的形式步骤。

到目前为止,坐标系的改变仅限于正交变换。从直角坐标过渡到斜角坐标的情况,或(ξ, η, τ)和(x, y, z)两者一开始就作为斜角平行坐标引入的情况,确实有待于补充。我们下面继续把坐标的原点看做为固定的。作出这个改变后,就从主群几何过渡到仿射群几何。当我们从仿射群几何来看代换系数在坐标变换下的情况时,发现代换系数尽管也表示着一个几何量的分量,但它们不像坐标乘积(1)一样变换,而是与之逆步变换。相应地,(6)式和(8)式中系数的情况也如此。可以证明,对于平行坐标系的同一个张量(例如,同一个均匀形变),可以用与量(1)同类的分量给出,也可以用(7)式中的系数那样的分量给出,前者称为张量的同步分量,后者称为张量的逆步分量。同步和逆步两个术语,常用"共变"和"逆变"来代替。有时,后两个说法在意义上是互换的。两类分量的区别,与点坐标和平面坐标之间的区别相同。

"张量"一词的另一个意义,以及比我们所取的意义更一般得多的意义,研究了齐次型在坐标改变下的情况后就会明白。前面,我们曾把这个研究转入到二次型

$$a_{11}\xi^2+2a_{12}\xi\eta+\cdots+a_{33}\tau^2$$

的情况。当然用的记号有点不同。我们发现二次型的系数 a_{11}, $2a_{12}$, …, a_{33}与点坐标的项成线性、齐次、逆步地变为 ξ^2, $\xi\eta$, …,

τ^2。但很容易看出，后者与(3)式成同步变换。我们可以把这个结果叙述如下：二次型的系数 a_{11}，$2a_{12}$，…，a_{33}是一个对称张量的逆步分量，项 ξ^2，$\xi\eta$，…，τ^2是对称张量的同步分量。相应的结果对双线性型也成立。关于双线性型，吉布斯认为，当可记作两个线性型的乘积时，我们可以说它构成一个并向量。最后，如果我们有一个点坐标的齐次 n 重线性型，我们只消稍微计算一下就可以证明，它们的系数在坐标变换下同样地作线性齐次变换，而且与点坐标项是逆步的。

上面所谈的张量概念的一般化，在于称每一个这样的量为张量，并不像我们以前那样，仅仅结合双线性型来用这个名称。爱因斯坦(Einstein)及其追随者就是在这个一般形式上用这个名称的。在旧一点的术语中，习惯上更常说线性型、二次型、双线性型、三线性型、三次型等。

除了这些术语不同外，在实践中还出现一种趋势，用单一的字母来记张量的分量组，并通过字母符号的组合来表示张量的运算。这一切符号本身实质上是非常简单的，如果读者觉得困难，只是因为不同的作者用不同的符号。这里所发生令人遗憾的情况与我们讨论矢量分析时提到的一样，但是这里的情况严重得多，看来不免有搞混的情况。因为全部现代文献中充斥着这种混乱的符号，所以我们不能不讲一讲。

现在让我们转过来谈度量几何，以便从中选几个有特征的例子。我要说明从不变量理论的系统讨论中导出两个重要的基本概念："两点 $x_1=\dfrac{\xi_1}{\tau_1}$，… 和 $x_2=\dfrac{\xi_2}{\tau_2}$，… 之间的距离 r"，以及"两个平面α_1，…，δ_1 和 α_2，…，δ_2 之间的角 ω"。从众所周知的解析几何公式得到

$$
\begin{aligned}
r&=\sqrt{(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2}\\
&=\sqrt{\frac{(\xi_1\tau_2-\tau_1\xi_2)^2+(\eta_1\tau_2-\tau_1\eta_2)^2+(\zeta_1\tau_2-\zeta_2\tau_1)^2}{\tau_1^2\tau_2^2}},
\end{aligned}
$$

$$\omega = \arccos\left(\frac{\alpha_1\alpha_2 + \beta_1\beta_2 + \gamma_1\gamma_2}{\sqrt{\alpha_1^2 + \beta_1^2 + \gamma_1^2}\sqrt{\alpha_2^2 + \beta_2^2 + \gamma_2^2}}\right),$$

它们分别是参数的代数函数和超越函数。如果我们证明构成它们的有理整式部分本身是在旧的意义下的不变量,则可以分别称它们为“代数”不变量和“超越”不变量。

我们从角 ω 出发,因该量的不变量式子由 $\alpha_1, \beta_1, \gamma_1, \delta_1$ 和 $\alpha_2, \beta_2, \gamma_2, \delta_2$ 的两个线性型和平面坐标的二次型

$$\alpha^2 + \beta^2 + \gamma^2 + 0 \cdot \delta^2$$

组成,后者代表了虚球面圆。我们当然能像前面用点坐标的形式构成不变量那样,通过交换点与平面坐标(“对偶化”),以这个平面坐标的二次型构造不变量,特别是对于两个给定值的组的二次型的值

$$\alpha_1^2 + \beta_1^2 + \gamma_1^2 + 0 \cdot \delta_1^2 \quad 与 \quad \alpha_2^2 + \beta_2^2 + \gamma_2^2 + 0 \cdot \delta_2^2$$

和这两个值组构成的极型的值

$$\alpha_1\alpha_2 + \beta_1\beta_2 + \gamma_1\gamma_2 + 0 \cdot \delta_1\delta_2$$

都是不变量。$\cos\omega$ 正是用这些式子构成的。进而,$\cos\omega$ 是两个值组 $\alpha_1, \cdots, \delta_1$ 和 $\alpha_2, \cdots, \delta_2$ 的零次齐次式,像所给的二次型的系数 1, 1, 1, 0 一样。所以这个表达式在度量几何学中有一个独立的意义。事实上,在度量几何学中,存在着与单位选择无关的绝对的角度测量。这说明,我们的表达式是一个绝对不变量。

其次,关于距离 r。我们回忆起,我们曾用一个或两个平面的坐标对点坐标的二次型的行列式“加边”而构成一个点坐标二次型的不变量。如果我们完全用对偶方法,对二次型 $\alpha^2 + \beta^2 + \gamma^2 + 0 \cdot \delta^2$ 的行列式

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{vmatrix},$$

用给出的点的坐标 $\xi_1,\cdots,\tau_1$ 和 $\xi_2,\cdots,\tau_2$ 加边一次和二次，则同样可以得到由平面坐标和两个点的坐标的二次型组成的不变量，从这样得到的行列式，我们作出商

$$\frac{\begin{vmatrix} 1 & 0 & 0 & 0 & \xi_1 & \xi_2 \\ 0 & 1 & 0 & 0 & \eta_1 & \eta_2 \\ 0 & 0 & 1 & 0 & \zeta_1 & \zeta_2 \\ 0 & 0 & 0 & 0 & \tau_1 & \tau_2 \\ \xi_1 & \eta_1 & \zeta_1 & \tau_1 & 0 & 0 \\ \xi_2 & \eta_2 & \zeta_2 & \tau_2 & 0 & 0 \end{vmatrix}}{\left\{\begin{vmatrix} 1 & 0 & 0 & 0 & \xi_1 \\ 0 & 1 & 0 & 0 & \eta_1 \\ 0 & 0 & 1 & 0 & \zeta_1 \\ 0 & 0 & 0 & 0 & \tau_1 \\ \xi_1 & \eta_1 & \zeta_1 & \tau_1 & 0 \end{vmatrix} \cdot \begin{vmatrix} 1 & 0 & 0 & 0 & \xi_2 \\ 0 & 1 & 0 & 0 & \eta_2 \\ 0 & 0 & 1 & 0 & \zeta_2 \\ 0 & 0 & 0 & 0 & \tau_2 \\ \xi_2 & \eta_2 & \zeta_2 & \tau_2 & 0 \end{vmatrix}\right\}}。$$

如果展开这些行列式，则不难证明，这个商正是上面给出的 γ 的值，因而是不变量。和前面考虑过的仿射几何中的基本不变量一样，这个商是齐次的，对两点的坐标而言是零次齐次，但对给出的二次型的系数而言是－4 次齐次的。而且它不是绝对不变量，因为每个行列式的权为 2，故商的权为 2－4＝－2。这里有与前一节结尾所考虑的作法相对称的方法，并与我们由此得出的结论一样。于是，数值 r 在度量几何中没有直接的意义。事实上，只有在假设了一个任意单位长线段，即将这样的线段和基本的二次型一起加入到结构中之后，才能测量两点间的距离。只有构造出这里所考虑的类型的各表达式的商，才能得到度量几何中的绝对不变量。这些例子至少能向你们说明系统引入整有理不变量而使仿射与度量几何得到系统发展的概

况。我希望你们通过阅读扩大你们从上述教科书里得到的知识[①]。

我稍微再讲一个所谓三角形几何学的简单例子，这个例子在克莱布什—林德曼的新版讲义中作了详细的讲解[②]。随着时间的推移，特别是由于中学教师的工作，在三角形几何学方面已经形成了广大的封闭的领域。这些中学教师竭力研究了许多不平常的点、线、圆，而这些都可以结合三角形，如重心、高度、角的平分、内接、外切、费尔巴哈(Feuerbach)圆等等来确定。长期以来人们一直在努力，而且至今仍在努力发现的无数关系，轻而易举地归入了我们的完整系统。设给出了 3 个点

$$(\xi_1, \eta_1, \tau_1), (\xi_2, \eta_2, \tau_2), (\xi_3, \eta_3, \tau_3),$$

作为一个三角形的顶点。因为我们关心的是度量关系，我们加入两个虚圆点，它的线方程是 $\alpha^2+\beta^2=0$。我们可以简单地把它们的点坐标(1, i, 0)和(1, −i, 0)加进去(图 20.5)。于是，整个三角形几何学只不过是这 5 个点，即 5 个任意点的投影不变量理论(其中两个点，我们用特殊的术语表明)。这个说明对三角形几何学的系统结构特征作出了清晰的分析，否则是看不清的。

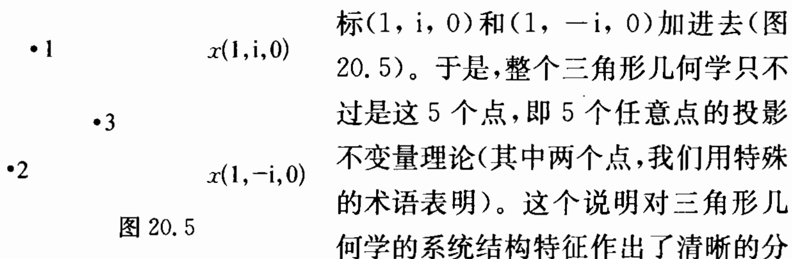


图 20.5

几何的系统发展的讨论，就到这里为止。上述严整的分类当然会使美感得到满足。不但如此，这个系统的处理本身可使我们对几何学产生更深的理解，所以每一个数学家和未来的教师都应该对它有所了解。基于这个理由，我觉得不得不把它包括在本课程内。尽

① 结合上述内容，应特别注意 H·伯克哈特在《数学年刊》第 43 卷(1893 年)里的文章："Über Funktionen von Vectorgrossen, welche selbst wieder Vectorgrossen sind. Eine Anwendung invariantentheoretischer Methoden auf eine Frage der mathematischen Physik"。

② 该讲义第 321 页。首先，我要提到《百科全书》(ⅢAB10)里的那篇文章。

管你们会在文献中经常找到这个观点，但也许不一定像这里所作的介绍那样系统。当然，把我们教条地束缚在这个系统处理上，总是从这个角度去讨论几何学，是完全不正确的。这样，很快就会变得乏味，失去一切吸引力。不管怎么说，会阻碍探索思想的发展，而探索思想始终是不受哪一种系统的束缚。

到此为止，我们所讨论的，在某种意义上是几何这座大厦的建筑设计。现在，我们的注意力将转到同样重要的一个问题，即几何的基础问题。

第二十一章

几何学基础

我们现在进入的是一个非常广阔的领域，这一章的观点取自F·恩里克斯刊于《百科全书》(Ⅲ′A. B. 1)的文章："Prinzipien der Geometrie"。

几何学基础的研究，往往非常接近认识论和心理学范围，因为认识论和心理学也要研究空间概念的起源及用数学方法来处理空间关系的合理性。当然，我们只能非常表面地触及这些问题，主要是处理涉及数学的一个侧面，并假定空间概念是想当然的事。至于个人的空间概念是怎样发展起来的，是怎样发展到我们数学家那样习以为常的精确形式的，这个问题在教育学上固然重要，但我们也必须略去不谈。

这样一限制，我们的问题就是：怎样在最简单的基础上，通过逻辑运算建立起整个几何学结构。纯粹逻辑当然不能提供基础。只有在问题的第一部分被解决后，即在有了由某些简单的基本概念与某些简单的陈述(所谓公理)组成的、符合我们最普通的直觉的系统之后，才可能用上逻辑推论。当然，这些公理可以根据著法的不同而被分为相互独立的各个部分，我们有很大的选择自由。这个公理系统所必须满足的一个条件是问题的第二部分所施加的：必须能从这些基本概念与公理出发，不求助于任何直觉，逻辑地推导出几何的全部内容。

本书全书结构就表明了这个问题的特定处理方法。原则上，我们始终利用分析的帮助，特别是利用解析几何方法的帮助。因此，这里也以一定的分析知识为前提。我们将问：怎样才能以最简捷的方法从一个给定的公理系统求出解析几何的定理？遗憾的是，由于几

何学家常常不愿使用分析，希望尽可能不用数字，因此很少使用这种简单的推导方法。

总之这里的研究一般地说可以用不同的方法来进行，取决于我们决定采用何种基本概念与公理。比较方便而且也是常用的方法，是从射影几何的基本概念，即从我们已经强调过的点、直线与平面出发。这些东西是什么？我们不想给出定义——我们一开始就必须知道这一点。我们只要对许多特征和相互关系作出陈述，以便由此按上面指定的意义推出几何学的全部内容。为此而需用到的公理现不完全列举，因为这样会使我们扯得太远。我只概括若干公理的内容，足以使你们获得一个清楚的概念。

首先是在投影几何中已说过的连接公理。开头不像投影几何里所要求的那样，毫无例外地要求平面上的两条直线必有交点或两平面必有交线。相反，为了适应于度量几何与仿射几何的关系，我们将限于下述公理：平面上两直线有一个公共点或没有公共点；两平面有一条公共线或没有公共线。然后再添加“不正常”的点、线与平面，从而导出射影几何的整个系统。

其次是顺序公理，说明平面与直线上不同点的相互位置。例如直线上 3 个点 a，b，c 必有一点例如 b 处在其他两点 a 与 c 之间，等等。这也叫做“介于性公理”(图 21.1)。

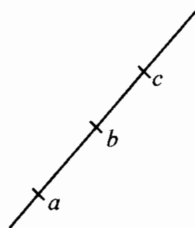


图 21.1

最后是关于连续性问题。暂时我只能强调直线上没有空隙。如果我们用任何方法将点 a 与点 b 之间的线段分割成 1 与 2 两部分(设 a 在 b 的左端)，使得部分 1 的所有点在部分 2 的所有点的左侧，则存在一分割点 c，使 1 的点位于 a 与 c 之间，2 的点位于 c 与 b 之间。这显然对应于用戴德金分割引入无理数的情况①。

① 见第一卷第二章第三节。

从这些公理出发，就能通过逻辑推导得出空间投影几何的全部内容。特别是能立即引入坐标，并用解析方法处理投影几何。

过渡到度量几何上来，则必须考虑在投影几何中还有∞^{15}共线群或空间投影变换的概念。我们知道如何把空间 7 个参数的运动主群描述为一个上述群的子群。它的不变量理论组成了度量几何学的基础。这个群由共线变换组成，并使某个平面即无穷远平面和该平面上的一条二次曲线即虚球面圆（或表示它的绝对极系统）保持不变。但是，如果我们希望正确地得到初等几何定理，则必须向前多跨一步。我们必须从主群中分离出 6 个参数的真运动子群（平移与旋转），它们与相似变换不同，保持两点间距离完全不变。用这种方法，我们将得到全等度量几何作为我们的不变量理论。例如：提出这样一个要求：一个运动的“路径曲线”就它只保持一个点固定而论是闭合的，就可以从主群中导出运动来。

这种建立几何学的构思或许从理论上来说是最简单的，因为它首先对投影几何只用到线性流形的运算，为了建立度量几何才不得不加上二次流形、虚球面圆。然而要实现这个构思，又是十分抽象和麻烦的事情，也许只有在专门讲投影几何的课程中才适宜于这样做。在这个一般的说明之后，就要请你们参考文献中最易读的表述，即 H·弗莱谢尔所译 F·恩里克斯的一本名为 *Vorlesungen über projective Geometrie* 的书①。

为了方便教学，对于马上就要讲到的这个几何课程，我宁愿采取另一种讲法。而且为了简单起见，我将以平面几何为限。

§21.1 侧重运动的平面几何体系

我们把点和直线作为基本概念，且对它们假设有连接、顺序和连

① 莱比锡，1903 年（德文版第二版出版于 1915 年）。原书书名为 *Lezioni di geometrie proieltiva*，意大利波洛尼亚，1898 年；第三版，1909 年。

续性公理。这里连接性公理仍然只包含这样的直观事实：过两点总可以连一条且只有一条直线，而两条直线则或只有一个交点或没有交点。关于在一条直线上的点的顺序，我保留前面已指明的条件。在研究的过程中，将会考虑对附加的顺序和连续性公理做仔细的阐述。

在这个基础上，我们将避免拐弯抹角地使用射影性，并立即引出平面上的运动群∞^3，并通过它而得出平面解析几何的系统。首先我们必须在一系列公理之内抽象地用公式表示将要使用的，关于我们的点线系统的这些运动的性质。当然，会以刚体作比拟，用以此得来的生动的运动概念作指导。于是，一个运动首先必须是空间的点的单值可逆变换。特别是，它必须使每个点与有限空间中的一个点是同等的。而且，它必须无例外地把一条直线变换成一条直线。一般来说，对这类变换，仍用"共线性"一词是方便的。当然，我们还不知道是否有这种共线性变换，因为现在我们不像以前那样掌握着投影几何。因此，必须通过一个新公理，至少对这些特别的共线性的存在作专门的假设。于是，我们假设存在一个我们称为运动群的共线性变换群∞^3，我们将把它的不变量理论作为平面几何学来看。我们必须较准确地解释这里的"无穷的三次方"表示什么。给定任两点 A, A'，分别用它们画出射线 a 与 a'（图 21.2），于是存在且仅存在一个把点 A 移到点 A' 且把射线 a 移到射线 a' 的运动。能借助运动把它们相互转移的图形，称之为全等。

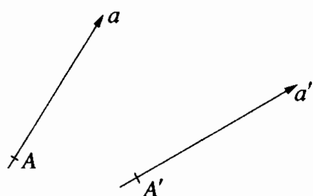


图 21.2

但我们不会使用这个运动群的全体，而只利用其中特殊的一类，我们将对此作某些专门的公设。事实上，正好只有一个运动把点 A 移到一个任意给定的点 A' 且将从 A 到 A' 的直线（连同这个方向）变成自己。我们称这样的运动为平移，或准确地说，称之为平行移动。我们现在假设，每个这样的平移将每一条连接任意对应点 B 与 B' 的

直线保持其正向而转换成其自身。而且关键的是，这个平面的∞^{12}个平移组成运动群的一个子群。

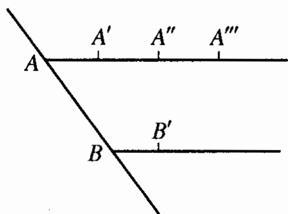


图 21.3

如果我们重复一个同样的平移(图21.3)，将点 A 沿着指向点 A' 的射线 AA' 逐次移到点 A', A'', A''', …。作为另外一个公设，我们必须假设这些点将最终到达或包括这条射线上的每个点。通过重复其逆变换，我们在反向射线上将得到一系列同类性质的点。如果我们设想每一个平移是从起点到终点连续地实现的(这一点以后我们会用到)，则将称所述的线为点 A 在平移下的轨线。因此，每条线都是无穷多个点的轨线，而对每个平移，有∞^1 个轨线，即转移为自身的直线。

现在应注意到，同一平移下的两条不同的轨线不可能相交。否则，此交点显然会是两个不同点平移的结果，即是从两条轨线中每一条得到的公共点，这与平移作为单值可逆变换的特点相矛盾。我们说，一个同样平移产生的所有轨线是彼此平行的。因此，从我们的运动的性质导出了这个概念。同时也很清楚，过一给定的点 A，必然存在一条与线 a 平行的直线，即在沿直线 a 平移的变换下的点 A 的轨线。

最后，我们必须为这些平移建立最后一个公理：任两个平移 T', T'' 是可交换的。即当我们对一给定点 A 首先施以平移 T' 然后施以 T''，其所得结果的点 B 与先施以 T'' 后施以 T' 的相同(图 21.4)，用符号可以写成 $T' \cdot T'' = T'' \cdot T'$。

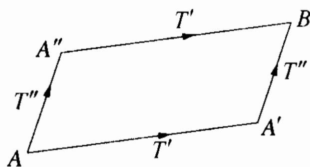


图 21.4

后面会讲到我们得出这些公理的方法。目前我们只强调，我们的原始公理只不过表达了每个人从一开始作几何图形时就已熟悉的

事情。几何作图中的第一件事就是移动一个固体,即从画图平面的一部分把直尺、圆规或其他仪器移动到另一部分以便转递一些度量。特别是经常沿一条直边滑动三角尺来实现平移的操作(图 21.5)。这些经验一次又一次表明,三角尺的所有点描出了平行线。因此,我们将不对这些假设作任何深入的逻辑分析,它们至少不是人为的。

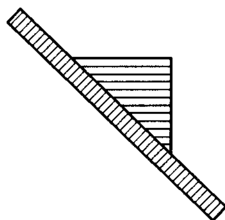


图 21.5

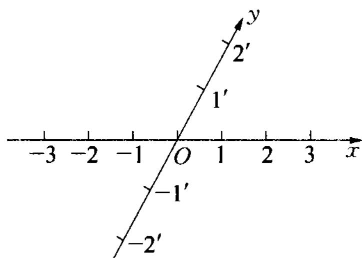


图 21.6

我们现在看看,用这些从平移中推导来的最初概念,能在解析几何中得到什么结果? 当然,我们不能谈论直角坐标系,因为我们还没有给直角下定义的基础,但可以引入一般的平行坐标系。我们过一点 O 画任意两条直线,分别称为 x 轴和 y 轴(图 21.6)。我们考虑把点 O 转移到 x 轴上任选的一点 1 的平移 T,并设重复平移 T 产生 x 轴上的点 2, 3, 4, …。如果用同样的方法进行逆变换 T^{-1},即通过平移把 1 变到零,则从点 O 逐次得到 x 轴上的点 −1, −2, −3, …。我们对这样得到的点指定正负整数 0, 1, 2, …, −1, −2, …作为“横坐标”x。确实它们并未包括 x 轴上的所有点,但根据我们的公设,它们的位置会使任何其他点均被包含在其中某一对点之间。

我们用类似的方法从沿 y 轴的任意平移 T' 出发,并向前或向后重复平移,得到点 1′, 2′, …, −1′, −2′, …,对它们指定相应的正负整数作为 y 的坐标,但这里应注意,我们对这样确定的 x 线段与 y 线段不能建立彼此的互换关系,因为尚未引入能使 x 轴变为 y 轴的运动(旋转)。

如果我们把任意确定的单位保持固定，那么现在就能考虑在 x 轴上具有非整数横坐标的点。我们首先讨论有理点。为了用一个例子把事情弄清楚，我们找一个沿 x 轴的平移 S，使得两次平移 S 就得到单位平移 T。我们对 O 应用平移 S 而得到的点指定为点 $\frac{1}{2}$。反复应用 S 即产生横坐标为 $\frac{3}{2}$，$\frac{5}{2}$，…的点。为了确定这种平移和这些点的存在，我们首先指出，从 $\frac{1}{2}$ 到 y 轴上点 $1'$ 的连线必平行于直线 $12'$（它对应于平分一个线段的已知作法）。事实上，如果我们把点 O 到点 $\frac{1}{2}$ 的平移 S（图 21.7）分解为由 O 到 $1'$ 的平移 T' 及由 $1'$ 到 $\frac{1}{2}$ 的平移 S'。按定义，两次平移 S 就等同于 T，根据两平移的可交换性，两次 S 就等于两次 T' 之后跟着两次平移 S'。因为两次 T' 将 O 变到 $2'$，所以等于是说，两次施用 S' 就会把 $2'$ 变成 1。于是 $2'1$ 是平移 S' 的轨线，与同样平移下的轨线的 $1'\frac{1}{2}$ 相平行。

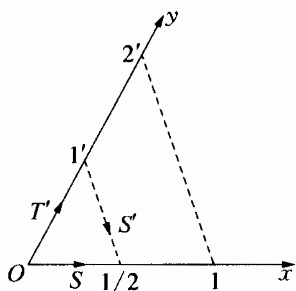


图 21.7

根据前述，我们已经有点 $2'$ 和点 1，从而有平移 S'。如果我们能证实从 $1'$ 出发的平移 S' 的轨线与 x 轴相交，则这个交点就是从给定的元素作出的点 $\frac{1}{2}$。而且这是唯一可能的。当然，从直观上没有人会怀疑这一点。但在我们的公理推导的结构中，我们需要一条专门的公理，所谓平面的“介于性公理”。根据此公理，如果一条直线通过一个边而进入一个三角形，则必然从另一边离开此三角形。这是空间直觉的普通事实，但由于它逻辑上独立于其他公理以外，因而需要这样强调。显然，完全类似的推理将证明每一个有理横坐标 x 均有对应的

点。从我们的公设不难推出，每一个线段无论怎样小，其内均存在“有理点”。

现在，为了真正达到在几何中实际考虑的所有点，我们必须引入无理数横坐标。为此需要一个新的并且又是十分明显的公理，对上述连续性的要求作出准确的陈述。如：在 x 轴上有无穷多个其他的点（轴到自身的平移），它们相对于有理点的顺序与连续性的关系，与无理数相对于有理数的顺序与连续性的关系一样。这个公理牵强之处在于：与此相反，从历史上来看，无理数是从几何的连续性引出的①。我们最后有：x 轴上的所有点与所有正负实数 x 具有单值可逆对应关系。类似的关系当然可以对 y 轴上的点建立起来。

我提醒你们注意，这样为在一条直线上建立刻度而设计的方法是完全自然的。当我们做刻度时，通常就是沿一条直的边界移动一个具有任意单位长度（例如说圆规两脚间的距离）的刚体，然后再细分这样得到的线段。

现在，平面上每一个沿 x 轴的平移，可以用一个简单的方程来表示，此方程对 x 轴上的每一个点 x 给出一个新位置的横坐标 $x'=x+a$。换句话说，有理或无理、正或负线段 a 被加到 x 上。类似地，一个沿 y 轴的平移用方程 $y'=y+b$ 表示。如果我们用任意顺序依次进行这两个平移，点 O 将变到一个确定的点 P（图 21.8），因为平移是可交换的。我们说 P 具有横坐标 a，纵坐标 b。反之，对任一点 P，可以指定唯一的两个数 a, b。我们只需将 O 移到 P，并决定坐标轴的交点相对于原来位置的横坐标与纵坐标。于是在平面上的点集与数对 (a, b) 集之间建

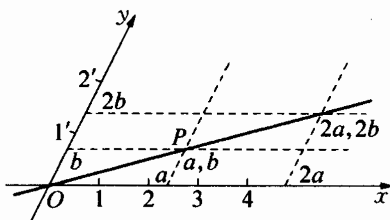


图 21.8

---

① 参考第一卷第二章第三节（无理数）。

立了一一对应关系，即在平面上有了一个完全确定的坐标系。

剩下的是考虑直线方程是什么样子的。我们首先研究从点 O 到点 $P(a, b)$的直线。它显然必须包含所有把点 O 变为点 P 的逐次平移所产生的所有点，即 $x=\lambda a$，$y=\lambda b$，其中 λ 是整数。而且我们看到，对有理数 λ，而后对无理数 λ，由这个方程所确定的点也在此直线上，但这已将直线上所有的点取尽了。消去 λ，我们得到直线方程 $x:y=a:b$，或 $bx-ay=0$。由此推出，只要 α，β 不同时为零，每个形如 $\alpha x+\beta y=0$ 的方程都代表过点 O 的一条直线。现在，任一条直线可由一个选好的通过点 O 的直线作平移而得出。于是，最终得出所有直线均由所有一次方程

$$\alpha x+\beta y+\gamma=0$$

给出。由于这个原因，它们被称为线性方程。

从直线具有线性方程这一事实，可推知大部分几何定理能毫无困难地用解析几何的方法推导出来，这里就不细说了，我只想补充说，我们可以用这种方法推导出全部仿射几何，因而可推出全部射影几何。在关于∞^2 个平移的子群的一个专门公设的基础上，我们能十分简单地得到这些。我只要再强调一下后面将要利用的一个事实。我们前面用射影几何的定理证明了莫比乌斯定理：每个共线变换是射影变换，即用坐标的线性分式函数或线性整代换给出的变换。现在，根据我们的第一假设，所有运动都是共线性的，且把每个有限点变为一个有限点。但另一方面，我们现在已推出了全部射影几何，因而从我们的角度出发，莫比乌斯定理是有效的。因此，每个运动将必然被一个前面引出的平行坐标 x 与 y 的线性整式变换所表示。

至此，我们还只能谈及 x 轴或 y 轴上两点之间的距离。如果我们希望深入到几何学的度量概念，特别是如果希望知道两条直线之间的角与任两点之间的距离，我们就必须把注意力转入整个运动群。

我们将专门考虑保持一个点，例如原点 O 不变的运动，即所谓

绕此点的旋转，根据关于一个运动的确定性的一般公设，正好存在一个把点 O 的射线 a 变到另一个任意的过点 O 的射线 a' 的旋转(图 21.9)。在某种意义上，旋转是平移的对偶变换，因为它们保持一个点不变，而平移把一条直线变为自身。正如平移一样，我们将把旋转考虑为从原来位置连续地进行的。因此，我们又要谈到每个点画出的轨线。

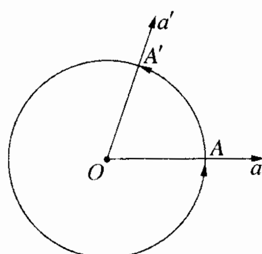


图 21.9

然而，在旋转和平移之间有一个本质区别，必须用一个公设来表达。通过反复施加一个同样的围绕点 O 的旋转，从 a 导出的射线 a', a'', …最终将达到或包含每一条过 O 的射线(而平移只产生一条射线上的点)。因此，射线 a 经连续旋转必然最终回到原来位置，而且 a 的每一点也回到原来位置。因此，轨线是一条封闭的曲线，它与每一条过点 O 的射线交于一点 A，从而所有线段 OA 彼此全部相同(即能通过运动把一个线段重叠于另一个之上)。这些轨线即通常所谓以点 O 为中心的圆。

借助这些旋转，我们将像用平移在直线上建立标度一样，在过点 O 的射线束上建立标度。这里，我们还得对连续性作出适当的假设。我不必详说，只给出结果：我们对每个旋转给一个实数，称为旋转角，且每个实数将作为一个旋转角而出现。当然，旋转的周期性将作为一个新概念而出现，自然会选把一条射线转成自身的完全旋转作为一个单位。但传统上，我们选一个完全旋转的 $\frac{1}{4}$ 作为单位，称其角为一个直角 R。于是每个旋转用角 $\omega \cdot R$ 来测量，其中 ω 可以是任意实数，但考虑到周期性，可能把角的值限制在从 0 到 4 上(图 21.10)。

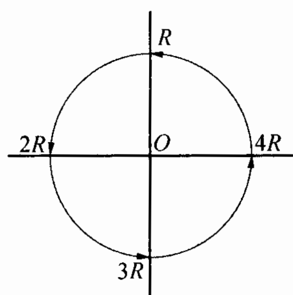


图 21.10

用同样的方法，可以确定出围绕任

意其他点 O_1 射线束的角的度数。但借助于平移，可以将点 O 处角的度数立即转移到点 O_1。事实上，如果给出了过点 O_1 的射线 a_1 与 a_1'（图 21.11），且如果 T 是将点 O 变到点 O_1 的平移，则平移之逆 T^{-1} 将 a_1 与 a_1' 变成过 O 的射线，记之为 a 与 a'。如果 Ω 是由 a 到 a' 的绕 O 的旋转，则由 a_1 到 a_1' 的绕 O_1 旋转 Ω_1，由 T^{-1}，Ω 和 T 逐次变换而成，或用符号写成

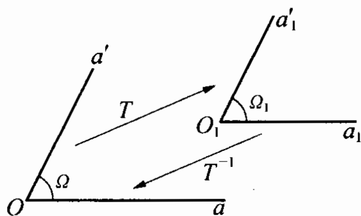


图 21.11

$$\Omega_1 = T^{-1}\Omega T。$$

这是由于：右侧也表示一个将 O_1a_1 变到 O_1a_1' 的运动，而这种运动是唯一的。现在，我们对 Ω_1 指定一个角 $\omega \cdot R$，而按上面定义，Ω 也具有同样一个角 $\omega \cdot R$。如果在线束 O 中有第二个旋转 Ω'，则在线束 O_1 中与之对应的旋转为

$$\Omega_1' = T^{-1}\Omega' T。$$

Ω_1 与 Ω_1' 的复合是

$$\Omega_1\Omega_1' = T^{-1}\Omega T T^{-1}\Omega' T = T^{-1}(\Omega\Omega')T,$$

它对应于 Ω 与 Ω' 的复合。由此得出，因在 O_1 处用上述旋转的复合而出现的度数实际上与原在 O 处的重复旋转而出现的度数相同。

在欧几里得几何中有一个被我们大多数初等教科书所删去的公理：所有直角是相等的。当然，每个中学生都会把这个公理看成是自然成立的。我也认为应该把它删去，因为学生不能理解它的意义。但它的内容与前面讨论的结果是相同的，即我们用在不同点的旋转所确定的等角，通过运动可以相重合。也就是说，它们是相等的。

我们现在已给出了角的一般定义，我们将定义任意两点之间的距离。至今我们只能通过平移来比较在同一条直线上的距离。如果

在 x 轴上取定了例如与 O 距离为 r 的点，则可以用绕 O 的旋转把它转移到过 O 的任意直线 a' 上(图 21.12)。于是可以把 x 轴上的长度标度转移到 a'，再通过平移而转到与 a' 平行的任何直线上，从而转到随便怎样的直线上。这样，通过用一直线连接两点，并用上述方法把 x 轴上的标度转到该直线上，就能测量任两点间的距离。特别是，我们将把开始为 y 所选的标度看成是这样从 x 轴上标度导出来的。

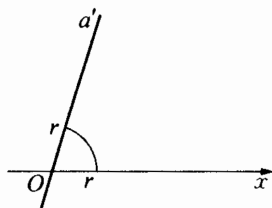


图 21.12

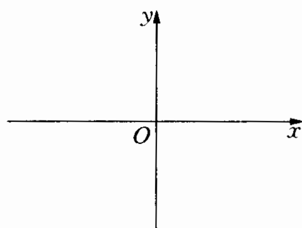


图 21.13

利用新的旋转概念，解析几何的工具就完备了。为此，我们将使用现在可以建立的直角坐标 x 和 y 来取代一般平行坐标系(图 21.13)。

我们已经知道，每个运动均由 x，y 的线性代换

$$x' = (a_1x + b_1y + c_1) : N,$$
$$y' = (a_2x + b_2y + c_2) : N$$

给出。因为它将每个有限点变为另一个有限点，分母 N 必然是常量且可以令其等于 1。如果专门考虑绕 O 的旋转，则 $c_1 = c_2 = 0$，我们有

$$x' = a_1x + b_1y, \qquad y' = a_2x + b_2y。 \tag{1}$$

对于绕过一个直角的特殊旋转，我们可以立即说出这个方程的形式。因我们有直角坐标系，x 轴变到 y 轴，y 轴则变到 $-x$ 轴，于是有

$$x' = -y, \; y' = x。 \tag{2}$$

这样，关于确定旋转公式的问题化为下面的纯分析问题：寻找一个形式为(1)的单无穷个代换的群，它包含代换(2)，且当 ω 是实参数时，群中的每个代换一般由(2)式经 ω 次迭代而产生。对 ω 为有理分数值，例如$\frac{p}{q}$，这种表示式当然表示重复 q 次后的代换给出迭代 p 次而得的代换(1)。对无理数 ω，则可按我们关于连续性的假设用有理数逼近。必须清楚地了解到，我们可以不预先假设具有任何几何知识，特别是关于直角坐标系旋转公式的知识。这样建立起来的结构，当然不能立即用于中学教学，但它确实具有非常简单和漂亮的形式。

我先说明，利用复数，代换(2)可用一个式子写成

$$x' + \mathrm{i}y' = \mathrm{i}(x + \mathrm{i}y)。\qquad (2')$$

从这个形式出发，我们看到两次施用此代换的结果，由 $x' + \mathrm{i}y' = \mathrm{i}^2(x + \mathrm{i}y)$ 表示。这是一个同样形式的方程，只是用因子 i^2 代替 i。类似地，在上述意义上的 ω 次重复，对每一个实数 ω 产生因子 i^ω。于是，我们对平面绕 O 经过角 $\omega \cdot R$ 的旋转得到其解析表达式为

$$x' + \mathrm{i}y' = \mathrm{i}^\omega(x + \mathrm{i}y)。\qquad (3)$$

为了准确地实现这个想法，我们必须使用分析中关于指数函数 e^z 以及三角函数的完整的知识，这些函数满足欧拉公式

$$\mathrm{e}^{\mathrm{i}z} = \cos z + \mathrm{i}\sin z。$$

在写出这个关系时，不必对其几何意义产生丝毫怀疑。

通过公式 $\mathrm{e}^{\mathrm{i}\pi} = -1$，我们也知道数 π，且我们可以写

$$\mathrm{i} = \mathrm{e}^{\mathrm{i}\cdot\frac{\pi}{2}}。$$

对 i^ω，我们在此理解为由公式

$$\mathrm{i}^\omega = \mathrm{e}^{\frac{\omega\mathrm{i}\pi}{2}} = \cos\frac{\omega\pi}{2} + \mathrm{i}\sin\frac{\omega\pi}{2}$$

所确定的唯一值。如果我们将之代入(3)式，并分离出实部与虚部，我们有

$$\begin{cases} x' = \cos\dfrac{\omega\pi}{2}\cdot x - \sin\dfrac{\omega\pi}{2}\cdot y, \\ y' = \sin\dfrac{\omega\pi}{2}\cdot x + \cos\dfrac{\omega\pi}{2}\cdot y。\end{cases} \tag{4}$$

它正是按初等解析符号所给出的运动群的表达式。

根据此结果，自然不取直角，而取$\frac{\pi}{2}$为单位。我们将像称自然对数一样，称它为自然角标度，以表明这些概念基于事物的自然性质，尽管要充分理解它的意义，尚需较深的洞察力。在这个自然标度中，我们用 ω 代替 $\omega\cdot\frac{\pi}{2}$，并得到代替(4)式的、作为旋转公式的著名方程

$$\begin{cases} x' = \cos\omega\cdot x - \sin\omega\cdot y, \\ y' = \sin\omega\cdot x + \cos\omega\cdot y。\end{cases} \tag{5}$$

现在必须考查这个公式，看看它们包含什么几何内容。它们原来是通常用来推出(5)式的一些初等定理。

(1) 我们从考虑 x 轴上到原点的距离为 r 的点 $x = r,\ y = 0$ 出发。如果将它转过角 ω，(5)式给出它的新坐标位置为

$$x = r\cos\omega,\ y = r\sin\omega。 \tag{6}$$

这里为了简单起见，省略了新坐标上的一撇。为了确定起见，设 $\omega < \frac{\pi}{2}$，并考虑由矢径 r 和点$(x,\ y)$的横坐标 x 与纵坐标 y 组成的直角三角形(图21.14)。(6)式表达了边与角 ω 之间的联系。根据三角函数的解析定义得出的关系式 $\cos^2\omega + \sin^2\omega = 1$，我们从(6)式立即得出

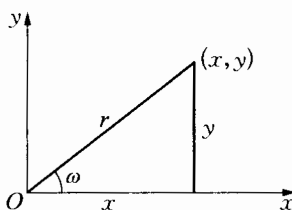


图 21.14

$$x^2 + y^2 = r^2。 \tag{6a}$$

这就是毕达哥拉斯定理①。在这里,它是我们关于平面上运动假设的结果。而且,我们可将(6)式写成

$$\cos\omega = \frac{x}{r}, \qquad \sin\omega = \frac{y}{r}, \tag{6b}$$

从而使我们的角函数得到了初等三角的意义。通常,角函数的严格定义形式是:余弦与正弦是邻边与对边分别和斜边之比。

(2) 如果我们将给出的元素、点或直线,通过平移或旋转变到前面考虑过的特殊位置,则不难为基本概念——距离与角给出一般的解析表达式。对于两点(x_1, y_1),(x_2, y_2)之间的距离,我们有

$$r = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}。$$

为了得到这个结果,只需要将点 2 用平移转到原点,根据平移公式,点 1 的新坐标就成了差:$x_1 - x_2$,$y_1 - y_2$,而(6a)式立即给出我们对 r 的表达式。同样,对方程为 $\alpha_1 x + \beta_1 y + \delta_1 = 0$,$\alpha_2 x + \beta_2 y + \delta_2 = 0$ 的两条直线之间的夹角 ω,可从(6b)式得到公式

$$\cos\omega = \frac{\alpha_1\alpha_2 + \beta_1\beta_2}{\sqrt{\alpha_1^2 + \beta_1^2}\sqrt{\alpha_2^2 + \beta_2^2}},$$

$$\sin\omega = \frac{\alpha_1\beta_2 - \alpha_2\beta_1}{\sqrt{\alpha_1^2 + \beta_1^2}\sqrt{\alpha_2^2 + \beta_2^2}}。$$

我不必给出证明的细节。

(3) 最后,我们尚需讨论面积的概念。到目前为止,在我们讲述几何中,还丝毫没有用到面积概念。然而,这个概念存在于每个人的自然空间意识中,即使不太严格。每个农民都知道一块地有多少亩是什么意思。尽管我们已完整地奠定了几何的基础而没有用到这个

① 我国称为勾股定理。——中译者

基本概念，但我们现在应该把它作为一个补充加入到系统以内，即用坐标来表达面积。

我们必须先进行一个简单的几何讨论，如同在欧几里得几何里进行的或在初等数学里进行的那种讨论。如果我们有一个边长为 A 与 B 的矩形，我们定义它的面积为乘积 AB。如果我们合并两个矩形或任意两个已知面积的图形，我们就定义合成图形的面积为两面积之和。如果我们从一个矩形或另一个图形中除去其中的一小块，则剩余部分的面积为原给两面积之差(图 21.15)。

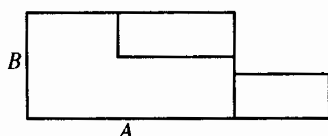


图 21.15

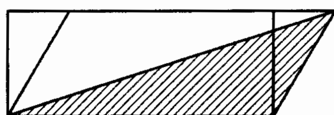


图 21.16

有了这些规定，我们立即可讨论平行四边形的面积。从一个底与高与其相等的矩形中去掉一个三角形，再加上一个与它全等的三角形，就得到这个平行四边形(图 21.16)。因此它的面积等于矩形的面积，即底与高之积。用一条对角线将平行四边形分成两个全等的三角形，因而其中每一个有平行四边形的一半面积，即三角形的面积是底乘以高之半。

如果我们将这个公式用于边长为 r_1，r_2，夹角为 ω 的三角形，边 r_1 上的高为 $r_2\sin\omega$，故面积为

$$\Delta=\frac{r_1r_2\sin\omega}{2}。$$

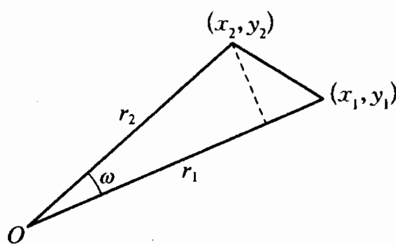


图 21.17

如果我们把这个三角形的一个顶点放在原点(图 21.17)，设另两点的坐标为 (x_1, y_1) 与 (x_2, y_2)，则根据上面得到的距离与角的公式，面积公式可写成

$$\Delta=\frac{x_1y_2-x_2y_1}{2}。$$

不难证明，坐标系的旋转使此公式保持不变，故它实际上提供了一个“几何概念”。为了得到在坐标平移从而在所有运动下的不变性，我们必须变换第三个顶点，即必须为具有 3 个任意顶点 (x_1, y_1)，(x_2, y_2)，(x_3, y_3)的三角形建立面积公式。我们用这种方法得到公式

$$\Delta=\frac{1}{2}\begin{vmatrix}x_1 & y_1 & 1\\ x_2 & y_2 & 1\\ x_3 & y_3 & 1\end{vmatrix}。$$

这事实上就是本卷开头的那个公式。不难证明，如果几个三角形被组合在一起或分成几个部分，则其面积按上面的公式相加或相减。正如我们以前所看到的，其证明依赖于简单的行列式的关系。

面积概念就这样补充进了我们的解析几何系统。同时我们获得了某些不包含在上述初步概念里的东西：面积变成了一个有符号的量。我在本卷的开头已讨论过，和面积作为一个绝对量的自然概念相比较，加上符号后以公式任意作运算，既通用又方便。

(4) 另外一个以不同程度的严格性在各个空间直觉中出现的概念，是一条(任意)曲线的概念。每个人都认为他知道曲线是什么；学了很多数学之后，才觉得有无数的怪事使他糊涂。这里我只简单地说，对我们而言，一条曲线是其坐标满足参数 t 的连续并可微分到所必要阶数的函数 ϕ, χ

$$x=\phi(t)\qquad y=\chi(t)$$

的点的全体。用这个方法，我们在解析几何的基础上，立即推导出通常称为微分几何的概念和定理，即弧长、曲面面积、曲率、渐屈线等等概念和定理。基本思想总是把曲线想象成折线的极限

(图 21.18)。如果两组邻点的坐标是(x, y)与$(x+\mathrm{d}x, y+\mathrm{d}y)$,则从毕达哥拉斯公式立即推出弧长公式为

$$\int\sqrt{\mathrm{d}x^2+\mathrm{d}y^2}。$$

以同样的方法,也可从以 O 为顶点的三角形面积公式推出曲线与两端点矢径间的扇形面积公式为

$$\frac{1}{2}\int(x\mathrm{d}y-y\mathrm{d}x)。$$

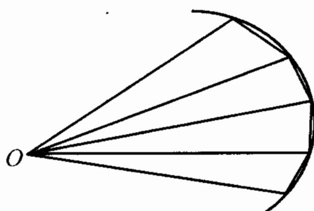


图 21.18

几何学的第一个发展体系就讲完了,这个体系的特点是首先确立了 3 个参变量的运动群的存在与结构,然后引入坐标,从而使我们后面的推论完全算术化。还有几何学的第二个发展体系,在某种意义上和上述体系正好相反,它直接导出度量几何,而且它始终占有一个重要地位。现在就来讨论这个体系。

§21.2 度量几何的另一种发展体系——平行公理的作用

和前一种发展体系相比,不同之处在于现在完全不要运动的概念,至多是把它作为事后考虑的对象。古代,甚至现代也常常这样进行讨论的一部分原因,在于哲学上的考虑,这一点我至少应该提到人们担心运动概念会把一个外部元素即时间概念引入几何学。后来试图用明显的刚体概念来说明引入运动概念的合理性时,又有人提出反对,认为这个概念本身不具有准确的可理解的意义。相反,人们主张只有掌握了距离概念之后,运动概念对我们才有意义。当然,经验主义者会回答说,抽象的距离概念,实际上只能从“足够”硬的刚体的存在中推出。现在对几何学的第二种几何发展体系的主要思想作一简略的介绍。

(1) 和以前一样，我们先引入点和直线以及关于它们的连接性、顺序性与连续性的公理。

(2) 除此之外，我们另外假设两个新的基本概念：一是两点之间(线段)的距离；二是两线之间的角。然后建立关于它们的公理，实质上线段和角可按习惯的方法用数来度量。

(3) 于是，第一个全等定理视为下述的特征性公理，它实际上取代了运动群公理：如果两个三角形有两边及其夹角分别相等，则它们是全等的，即它们的各部分均相等。在上一个体系里，这是一个可证明的定理，因为我们可以找到一个运动，把边 $A'B'$ 移到与 AB 重合(图 21.19)。于是，$A'C'$ 必然与 AC 重合，两个三角形也由于这个假设而重合了。但如果不把运动作为基本概念包括在内，即如果不使用运动概念，就不可能证明此定理，就必须假设它为一个新公理。

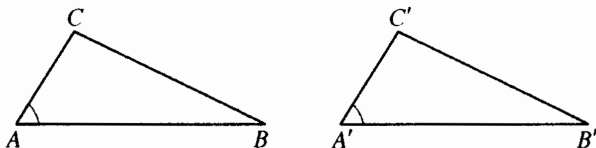


图 21.19

(4) 其余步骤，正如你们所知道的，和上一个体系完全相反。初等几何教学都是这样做的，本质上遵循着后面还要谈到的欧几里得体系。习惯上首先是证明毕达哥拉斯定理，然后从它们在三角理论中的意义引出三角函数余弦与正弦。从这里开始，最后导出我们已介绍过的同样的解析工具。

(5) 在这个过程中，建立另一个十分重要的关于平行理论的公理就成为必要了。在上一个体系中，平行性是随平移的考虑之后立即出现的最初几个基本概念之一。那里说，几条直线如果是同一平移的轨线，就称为平行的。这里则完全不同：平行性还不属于已经考虑过的基本概念之列，所以现在必须加以讨论。事实上，如果我们

有一条直线 g（图 21.20）和其外一点 O，将 O 与 g 上一点 P 连接起来，并使 P 沿 g 移动而通过点 P'，P''，…。换句话说，我们考虑一个个点 P，P'，P''，…，或一条条直线 OP，OP'，OP''，…。这里没有原先意义上的运动思想。射线 OP 在 P 移到无穷远时达到一个极限位置，我们称这个极限的直线为通过点 O 与 g 平行。当 P 从两个方向趋向无穷时，并未表明 OP 应趋向同一极限位置，以致有这样一种抽象的可能性：过点 O 存在两条与 g 平行的不同的直线。

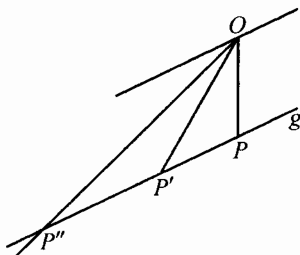


图 21.20

因此，在现在这个体系中，如果根据我们的一般直觉，我们假设两个极限位置应一致，即过一点只有一条直线平行于给定直线，那就是一条新公理。关于这一条著名的平行公理争论了好几百年，它也称为欧几里得公理，因为它是欧几里得作为一个公设提出的。

我想向你们介绍一下这个公理的历史。多少年来，人们尽最大努力企图证明这个公理，即证明它是其他公理的推论，但都失败了。当然即使到今天，还没有完全放弃这种企图。因为，尽管科学进展到现在，总会有人自以为是，对经过严格研究而得出的确凿结果一无所知。事实是：数学的进展早已超越这些徒劳无益的企图而进入了有成效的新的研究，并取得了积极的成果。早在 18 世纪，就有人提出过以下典型问题，暗示着有一条新的路子：如果把平行公理放到一边，是否可能建立一个没有矛盾的、逻辑上协调一致的几何系统，允许两条不同的极限直线以上面所讨论的意义而存在，即过点 O 允许作 g 的两条不同平行线。

到 19 世纪初，这个问题已得到确定的回答。高斯第一个发现了“非欧几何”的存在，而非欧几何就是他对这种几何体系的命名。从他死后发表的论文中看出，他在 1816 年已肯定了解了这一点。不过

他的研究笔记很久以后才被发现，直到1900年收到他的选集的第8卷[①]。高斯本人除了偶然提过外，对这个伟大的发现未曾发表什么见解。法律学家施韦卡特(Schweikart)也在1818年左右独立于高斯创立了非欧几何学，他称之为星球几何学，但同样未发表结果，这是从后来被发现的一封给高斯的信中得知的。最初发表非欧几何的人是俄国的罗巴切夫斯基(Lobatschefsky，1828年)和匈牙利的J•B•鲍耶(J. Bolyai de Bolya，1832年)，他们彼此独立地得到了这些结果，并分别于1826年和1823年得到了证明[②]。在19世纪，通过大量的论文，科学家们已普遍掌握了这些结果。今天，每一个有文化的人都已听说过非欧几何，尽管只有专家才有清楚的了解。

在19世纪后半叶的初期，黎曼对这些问题提出了一个全新的方向。1854年，黎曼在他的住所以“Über die Hypothesen welche der Geometrie zugrunde liegen”为题发表的演说[③]中指出，前人的一切研究，都以直线是无限长的假设为基础。这当然是十分自然和明显的。但他问：如果我们取消这个假设，即我们允许直线像地球上的大圆一样返回到自身，那又是什么样的情况呢？我们就会面临空间的无穷性与无边界性之间的差别。这一点，或许可以在二维空间中很好地看到。球的表面和普通平面都没有边界，只是后者是无穷的，而前者是在有限范围内的。黎曼事实上假设空间仅仅是无边界而不是无穷的。于是，有许多点位于其上的直线，会像一个圆一样，是一条封闭曲线。如果我们令一点 P 像前面一样在一条直线 g 上按一

① 莱比锡，1900年。所涉及的那一部分是P•施塔克尔(P. Stäckel)编的。

② 由恩格尔(Engel)和施塔克尔两人译成德文，发表于*Urkunden zur Geschichte der nichteuklidischen Geometrie*一书中。该书第一部分(关于罗巴切夫斯基)由恩格尔译(莱比锡，1898年)；第二部分(关于鲍耶)由施塔克尔译(莱比锡，1913年)。还可参阅斯塔克尔和恩格尔著*Urkundensammlung zur Uorgeschichte der nichteuktidischen Geometrie*，莱比锡，1895年。

③ 发表于*Abhandlungen der Gesellschaft der Wissenschaften zu Göttingen*第13卷，或参阅其《数学著作集》，第二版(莱比锡，1892年)，第272页及以后部分(第三版由柏林施普林格出版社于1923年出版，H•外尔(H. Weyl)译)。

确定方向越移越远，它最终将回到其原来的位置。前面讨论的射线 OP 将不会有一个极限位置，同时不会有过点 O 平行于 g 的直线。因此，与高斯等人的非欧几何相比，出现了属于黎曼的第二类非欧几何。

初看之下，这似乎是怪论，但数学家们立即注意到两种非欧几何与普通的二次方程理论的关系，也对了解这个问题提出了一个方法。确实，一个二次方程，或有两个不同的实根，或没有实根（两个虚根），或作为一个过渡情形，把一个实根计算两次。这完全和高斯的两条不同的平行线，黎曼的不存在平行线，以及介乎两者之间的过渡情况——一条平行线作为欧氏几何中同样的极限位置而计算两次的情况相类似。

在我比较仔细地讨论非欧几何之前，我至少要简单地讲一下它的极大的哲学意义，正因为如此，它始终引起哲学家的巨大兴趣，但也往往遭到他们的断然否定。

首先，这个新的分支为从纯逻辑观点去看几何公理的性质铺平了道路。由于非欧几何的存在，我们立即能得出结论：欧氏公理不是先行的基本概念和定理的推导的结果，我们也没有接受它的任何逻辑必然性。因为如果我们保留所有其他的公理，但用一个相反的假设去代替这条公理，我们并不会导致矛盾；相反，我们可以得出像欧氏几何一样正确的逻辑结构——非欧几何。诸如平行公理所描述的我们的空间图形感觉，当然不是一种纯逻辑的必然。

这样，问题就是：我们或许是否能借助于感性知觉来断定平行公理的正确性。对于这一点，非欧几何也作了明确的阐述。确实，感性知觉当然不会告诉我们只存在一条平行线，因为我们的空间知觉肯定不是绝对准确的。像在其他每一个感性知觉范围中一样，这里我们也不能分辨差别在某一限度以下的，即所谓知觉阈以外的度量（线段、角度等）。例如，如果我们从点 O 画出彼此十分靠近的两条直线（图 21.21），使它们的交角充分小，例如 $1''$ 或 $0.001''$，甚至更小，

我们肯定不再能区分它们。因此，通过感性知觉很难断定，是否真的只存在一条过点 O 的 g 的平行线，抑或存在两条相互分离角度十分小的平行线。如果我们设想点 O 离 g 十分远，例如远至天狼星或更远百万倍，我们就能更清晰地意识到这一点。在这样的距离下，感性知觉就完全失去了我们所期望的敏锐性，我们肯定不再能凭视觉来断定，旋转射线的极限位置是提供了一条还是两条平行于给定直线 g 的直线。

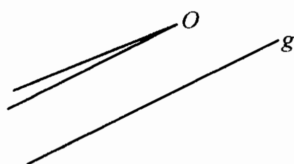


图 21.21

这种情况实际上既与欧氏几何一致，又与前一类非欧几何一致。我们接着就要看到，当我们深入考察数学公式时，发觉其中总包含着一个任意的常数。通过对此常数的适当选择，只要 O 到 g 的距离适当远，就可使两条平行于 g 的直线的交角任意小；仅当 O 到 g 的距离十分大时，这个角才大到可以分辨的程度。鉴于我们的空间直觉只适应于有限的空间部分，因此只能在有限的精确度以内，它才能明显地被前一类非欧几何如我们所希望那么近似地满足。

对于黎曼非欧几何来说，情况也类似。只是必须理解，直线的无限长度是不能从我们的感性知觉推导出来的。我们只能在有限的空间部分以内跟踪一直线，因而如果我们说此直线有非常大的但毕竟是有限大的长度，哪怕是比到天狼星的距离大百万倍以上，这也是与我们的空间经验不相矛盾的。想象力可以造出超过任何直接知觉可能性的任意大的长度。根据这些考虑，我们可以借助黎曼非欧几何所希望的精度表示空间的任何有限部分的情况，因为这样的几何学也有一个任意常数。

这里所触及的逻辑事实与直觉事实，是从数学的立场看的，与许多哲学家归在康德名下的那种空间观念有很大的冲突，因为根据康德的空间观念，一切数学定理都必须有绝对的效果。非欧几何自从被介绍到哲学界以来吸引了如此的关注并招致如此的反对，其原因

即在于此。

如果现在要我们去对非欧几何进行适当的数学解释的话，我们将尽可能选择通过射影几何来解释的道路。这就是我在 1871 年《数学年刊》第 4 卷中所给出的推导①。

正如在本章开始讨论几何学基础时所简短提示的，我们把射影几何看做是从点、线、平面这几个基本概念，以及从连接性公理、顺序公理及连续性公理发展起来的，与一切度量无关。特别是，我们引入点坐标 x, y, z 或齐次坐标 $\xi : \eta : \zeta : \tau$ 和平面坐标 $\alpha : \beta : \gamma : \delta$，使得点和平面的相互一致性由双线性方程

$$\alpha\xi + \beta\eta + \gamma\zeta + \delta\tau = 0$$

给出。

在此基础上，借助于凯莱原理和不变量理论，并引入用平面坐标写成

$$\phi_0 = \alpha^2 + \beta^2 + \gamma^2$$

的特殊二次型(令其等于零即表示虚球面圆)，我们已经建立起普通的欧氏几何。如我们已指出的，两平面间的角

$$\omega = \arccos \frac{\alpha_1\alpha_2 + \beta_1\beta_2 + \gamma_1\gamma_2}{\sqrt{\alpha_1^2 + \beta_1^2 + \gamma_1^2}\sqrt{\alpha_2^2 + \beta_2^2 + \gamma_2^2}}$$

和两点间的距离

$$r = \frac{\sqrt{(\xi_1\tau_2 - \xi_2\tau_1)^2 + (\eta_1\tau_2 - \eta_2\tau_1)^2 + (\zeta_1\tau_2 - \zeta_2\tau_1)^2}}{\tau_1\tau_2}$$

同时是所给出图形(两平面或两点)和二次型 ϕ_0 的简单的不变量。

我们打算用同样的方法建立非欧几何。我们用另一个“接近”虚

① “Über die sogenannte nichteuklidische Geometrie”，第 573 页及以后部分，或参阅《数学著作集》第 1 卷第 254 页及以后部分。

球面圆 $\alpha^2+\beta^2+\gamma^2=0$ 的二次型,即

$$\phi=\alpha^2+\beta^2+\gamma^2-\varepsilon\cdot\delta^2$$

来代替前者,其中 ε 是我们可以选择的任意小的参数,且当 $\varepsilon=0$ 时,有 $\phi=\phi_0$。我们的二次型是这样选择的:从正 ε 得到第一类非欧几何,从负 ε 得到黎曼几何,而从 $\varepsilon=0$ 则得到普通欧氏几何的公式。作出这个型的关键在于:其行列式

$$\Delta=\begin{vmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-\varepsilon\end{vmatrix}=-\varepsilon$$

一般不为零。仅当特殊情况 $\varepsilon=0$ 时,即当 $\phi=0$ 代表虚球面圆时,此行列式才等于零。于是,我们的假设归结为:用一个行列式为正数或负数,但其绝对值可任意小的二次型,取代行列式等于零的二次型。

我们从一般的二次型 ϕ 和两平面或两点组成的图形建立起不变量,这些不变量完全类似于欧氏几何中表示特殊型 $\phi_0=\alpha^2+\beta^2+\gamma^2$ 的不变量,以求得非欧几何的度量系统的定义。这只不过是凯莱于1859年发展的概念①:"人们可以把虚球面圆完全看做是任意二次曲面(例如曲面 $\phi=0$),同样地定义出一套度量系统。"由于篇幅有限,权且事先建立解析公式。这样就可用精确的形式最快地概述其内容,避免任何神秘的色彩。当然,只有以后从几何方面去仔细研究,就像我在上述《数学年刊》第4卷的文章中所做的那样,才能由这个说明进而达到对内容的完全理解。

我们首先考虑两个平面。把前面的交角表达式推广,很自然地建立两平面之间相对于曲面 $\phi=0$ 的交角的表达式。正和前面一

---

① 见前面已经引过的《关于四元数的第六个专论》一文,见本章第一节。

样，我们从二次型 ϕ 和它的极型的值，建立公式

$$\omega = \arccos \frac{\alpha_1\alpha_2 + \beta_1\beta_2 + \gamma_1\gamma_2 - \varepsilon\delta_1\delta_2}{\sqrt{\alpha_1^2 + \beta_1^2 + \gamma_1^2 - \varepsilon\delta_1^2}\sqrt{\alpha_2^2 + \beta_2^2 + \gamma_2^2 - \varepsilon\delta_2^2}}。$$

用这个方法得到一个表达式，它显然是不变量，且 $\varepsilon = 0$ 时即为欧氏几何的交角公式。

把两点间距离的表达式转换到非欧度量系统，就不是那么直截了当。困难在于，现在有一个行列式不为零的二次型代替了行列式为零的二次型 ϕ（欧氏度量系统的特征）。但是，如果我们用严格的对偶方式得出交角的定义，我们就能发现怎样去建立距离的表达式。这样，我们就一定能得到一个不变量。首先，我们建立点坐标的曲面方程 $\phi = 0$。前面讲过，用点坐标对 ϕ 的行列式 Δ 加边，我们得它的左侧 $f(\xi, \eta, \zeta, \tau)$

$$f = \begin{vmatrix} 1 & 0 & 0 & 0 & \xi \\ 0 & 1 & 0 & 0 & \eta \\ 0 & 0 & 1 & 0 & \zeta \\ 0 & 0 & 0 & -\varepsilon & \tau \\ \xi & \eta & \zeta & \tau & 0 \end{vmatrix} = \varepsilon(\xi^2 + \eta^2 + \zeta^2) - \tau^2。$$

现在写出 f 的极形式，除以 f 在点 1 和点 2 之值的平方根的积，然后取余弦，把 ω 的表达式转变为

$$\gamma = k\arccos \frac{\varepsilon(\xi_1\xi_2 + \eta_1\eta_2 + \zeta_1\zeta_2) - \tau_1\tau_2}{\sqrt{\varepsilon(\xi_1^2 + \eta_1^2 + \zeta_1^2) - \tau_1^2} \cdot \sqrt{\varepsilon(\xi_2^2 + \eta_2^2 + \zeta_2^2) - \tau_2^2}}。$$

在此插入因子 k，以便使我们能按习惯做法那样选择任意线段为单位长。而且，当我们转回到欧氏几何上时，这也是必要的。我们必须设想，当 ε 是负数时，k 是实的；当 ε 是正数时，k 是纯虚数，以便使得对于所有实点，或至少在所有实点的某个子区域内（当 $\varepsilon > 0$ 时），r 都是实数，从而为非欧几何给出实的基础。

我们现在已得到距离的一般定义。剩下只需证明，当 $\epsilon=0$ 时，得出欧氏几何中通常的表达式。这不像对角 ω 所作的那样容易。因为如果令 $\epsilon=0$，则其商即为 1，而 $\frac{r}{k}$ 等于零或任意 2π 倍数。尽管有这样一个奇怪结果，我们还是可以通过某种办法最终求得欧氏表达式。方便的办法是：通过关系 $\arccos\alpha=\arcsin\sqrt{1-\alpha^2}$，把定义 r 的方程改变。消去一个公分母之后，我们得到 r 的值为

$$k\arcsin\sqrt{\frac{\{\epsilon(\xi_1^2+\eta_1^2+\zeta_1^2)-\tau_1^2\}\{\epsilon(\xi_2^2+\eta_2^2+\zeta_2^2)-\tau_2^2\}-\{\epsilon(\xi_1\xi_2+\eta_1\eta_2+\zeta_1\zeta_2)-\tau_1\tau_2\}^2}{\{\epsilon(\xi_1^2+\eta_1^2+\zeta_1^2)-\tau_1^2\}\{\epsilon(\xi_2^2+\eta_2^2+\zeta_2^2)-\tau_2^2\}}}。$$

我们现在可以很容易地改变分母。事实上，可使用已知的行列式关系证明，f 在点 1 的值（即二次型 Φ 的被加了一次边的行列式 Δ）乘以在点 2 的同样行列式，减去对点 1 与点 2 取的极形式，等于行列式 Δ 与将它用点 1 与 2 的坐标加边两次所得的行列式之积，即等于积

$$\begin{vmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-\epsilon\end{vmatrix}\cdot\begin{vmatrix}1&0&0&0&\xi_1&\xi_2\\0&1&0&0&\eta_1&\eta_2\\0&0&1&0&\zeta_1&\zeta_2\\0&0&0&-\epsilon&\tau_1&\tau_2\\\xi_1&\eta_1&\zeta_1&\tau_1&0&0\\\xi_2&\eta_2&\zeta_2&\tau_2&0&0\end{vmatrix}。$$

完成这个乘法后，可得

$$\begin{aligned}-\epsilon\cdot\{&(\xi_1\tau_2-\xi_2\tau_1)^2+(\eta_1\tau_2-\eta_2\tau_1)^2+(\zeta_1\tau_2-\zeta_2\tau_1)^2\\&-\epsilon(\eta_1\zeta_2-\eta_2\zeta_1)^2-\epsilon(\zeta_1\xi_2-\zeta_2\xi_1)^2-\epsilon(\xi_1\eta_2-\xi_2\eta_1)^2\}。\end{aligned}$$

任何一个不熟悉行列式计算技巧的人都可以用直接变换证明，这个表达式恒等于前面 r 的表达式的分子。如果将此式子插入 r 的公式，并令 $\epsilon=0$，由于因子$\sqrt{-\epsilon}$，当然得到和第一个形式一样的结果

$$\frac{r}{k}=\arcsin 0=0。$$

但如果我们不允许 $\varepsilon=0$，只使它变得非常小，作为一级近似，arcsin 将等于 sin。在分子中可以忽略各乘以 ε 的 3 个平方项，而在分母中可以忽略乘以 ε 的项。作为一级近似，剩下了

$$r=k\cdot\sqrt{-\varepsilon}\frac{\sqrt{(\xi_1\tau_2-\xi_2\tau_1)^2+(\eta_1\tau_2-\eta_2\tau_1)^2+(\zeta_1\tau_2-\zeta_2\tau_1)^2}}{\tau_1\cdot\tau_2}。$$

我们现在已找到了前面说过的方法。在过渡到极限 $\varepsilon\to 0$ 的过程中，我们对 k 不指定一个固定值，而是使其变成无穷大，并使 $\lim(k\cdot\sqrt{-\varepsilon})=1$。为此，当然必须根据 ε 趋向零时是取正值或负值而分别使 k 取纯虚数或实数。但在这个过渡到极限的过程中，显然可取得欧氏几何的距离表达式。

按照我们的方式去考虑二次型 f 的几何意义以及这里的解析表达式的意义，就可以得出结论：实际上，$\varepsilon>0$ 是第一类非欧几何的情况，$\varepsilon<0$ 是第二类非欧几何的情况，而 $\varepsilon=0$ 当然是欧氏几何的情况。这里不能给出全部论据，为此必须请你们参考我在《数学年刊》第 4 卷中的文章①。那个时候，我曾建议把这 3 类几何定名为双曲形、椭圆形和抛物形，因为两条实、两条虚或两条重合的平行线的存在，恰好对应于 3 种圆锥曲线的渐近线的情况，以后你们会在文献中经常看到这些名称。

我将用例子详细说明，从距离的表达式中会得出怎样的平行线定理。为此在平面上选择双曲型几何的例子。于是必须令第三个坐标等于零。我们的二次型变成 $\Phi=\alpha^2+\beta^2-\varepsilon\delta^2$，当它等于零时代表一条实圆锥曲线，且因 $\varepsilon>0$，我们可以设想它为一个椭圆。距离公式取形式

---

① 又需注意我写的“Einführung in die nichteuklidische Geometrie”，这是我的非欧几何讲义油印本的修订本，即将出版(W·罗森鲍姆(W. Rosenbaum)编)。

$$\gamma = k\arccos\frac{\varepsilon(\xi_1\xi_2+\eta_1\eta_2)-\tau_1\tau_2}{\sqrt{\varepsilon(\xi_1^2+\eta_1^2)-\tau_1^2}\sqrt{\varepsilon(\xi_2^2+\eta_2^2)-\tau_2^2}},$$

其中 k 为纯虚数。不难理解，对于处在此圆锥曲线内部的点，这个公式取实数值，我们把内部的点理解为平面内的所有点，从这些点作不出圆锥曲线的实切线。因此，双曲型几何学的运算域完全由这些内点和位于内部的直线组成。圆锥曲线上的点(图 21.22)本身表示无穷远域，因为由公式得到的每个点 1 到圆锥曲线上点 2(对它，$\varepsilon(\xi_2^2+\eta_2^2)-\tau_2^2=0$) 的距离为$\infty$。因此，在这个意义上，在双曲型几何中，每条直线上有两个无穷远点，即为它与圆锥曲线$\Phi=0$的交点。但每条射线 a 只有一个无穷远点。如果有一条直线 g 和不在其上的一点 O，则按前面的定义，过 O 的平行线是连接点 O 与 g 上移向无穷远的点之连线的极限位置，而在这里就是连接点 O 与 g 和圆锥曲线交点的直线。于是，有两条本质不同的平行线，每一条附属于 g 的两个方向中的一个。

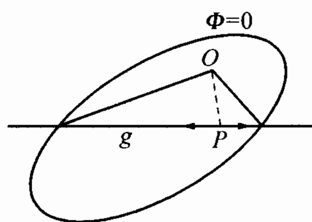


图 21.22

请再同上述欧氏几何第一个体系作一番比较。该体系从运动群出发，而运动群就是保持度量不改变的全部共线变换。在非欧几何中有类似的这种共线变换。一个一般的二次齐次方程有 10 项，故有 9 个本质常数。在最一般的空间共线变换中，共有 15 个任意参数，故存在一个六重无穷的共线变换，它把给定的二次型(例如二次型 Φ)变成自身。确实，这是使我们导出的度量关系保持不变的条件。因此，在每种非欧几何中也存在一个六重无穷的“运动”群，它保持 ω 和 r 不变。对于平面上的情况，参数会像以前一样减为 3 个。

因此，我们也能从一个运动群的存在出发来发展每种非欧几何。余下只需指出，较早的体系怎样会单单导出欧氏几何。当然，其理由在于：我们从运动中选出了特殊的所谓平行移动的两参数(在空间

中是 3 个参数)子群,它的轨线都是直线。在任何非欧几何中没有这样的子群,又因我们一开始就假设它们是存在的,所以就永远不可能导出非欧几何,而只能导出欧氏几何。

最后,让我提出若干建议,以结束对非欧几何的专题讨论。

(1) 我上面已经讲过,在哲学界,以往非欧几何经常不能得到充分的理解。但另一方面,我必须强调的是,今天它已被数学科学所广泛承认。事实上,它有许多用途,例如在现代函数论和群论中,它已被当作十分方便的手段,使算术上复杂的种种关系变得一目了然。

(2) 每一个教师都肯定应该懂点非欧几何,它已经是广泛谈论的少数几个数学部分之一,至少有的人常常把它挂在嘴边上。所以,任何教师都会随时随地被问到这方面的问题。在物理学上,像这样人人挂在嘴边的问题当然更多,每一个教师不可不知。事实上,物理学上几乎每一项发明都属于教师应知的范围。身为物理教师,却说不上什么叫伦琴射线,什么叫镭,这是能够想象的吗? 一个回答不了非欧几何问题的数学教师,也不会给人以很好的印象。

(3) 另一方面,我又着重建议,不可像某些热心人始终建议的那样,把非欧几何引入正规的中学课程(除非感兴趣的学生问到的时候偶尔作一些提示)。只要遵照上述建议,让学生真正学会欧氏几何,我们也应该心满意足了。不过无论如何,教师总应该比一般学生多知道一些。

我想再简单地讨论一下非欧几何所引起的现代数学的进一步发展。非欧几何的一个结果,即欧氏平行公理在逻辑上是独立于其他公理的,已经造成了一个良好的开端。它推动了对其他几何公理之间的逻辑依赖性或独立性的研究。由此产生了现代几何公理理论,紧紧地遵循旧的研究已开辟的道路,确定了哪些几何部分可以不用某些公理而建立起来,也确定了是否可以通过与给定公理相反的假设而获得一个没有矛盾的体系,即所谓“拟几何学”。

这一方面最重要的著作,是希尔伯特的《几何基础》[①]。与以往的研究相比,这本书的主要目的是以上述方式确定连续性公理的意义。这当然首先要重新安排几何公理系统,使得一直放在开头的连续性公理在结尾出现。例如,在我们的非欧几何体系中,不能利用把平行概念放在开头的关于公理的第一种安排。相反,不得不另外建立一套公理系统,使大部分讨论不涉及平行性,而将平行公理加在最后。希尔伯特公理系统把上述重大的不同丢在一边,主要与我们讲的初等几何的第二种发展相一致。

在这个意义上可以说,希尔伯特对不用连续性公理可以把几何学发展到什么程度进行了探讨。他的讨论也把"拟几何学"包括在内,其中除了连续性公理外,所有其他几何公理都有效。这种几何本质上是直线上的点和普通实数(点的坐标)一一对应(见本章开头部分及第一节)。当然,我不可能详细介绍希尔伯特的研究思路,以及他所得到的,关于某些几何定理与公理之间逻辑联系的有趣的结果。我只作这点解释,关于上述一切,请你们去读希尔伯特的原著。但我要提醒的是,我们在上卷中讨论过的内容就是他的非阿基米德几何[②],而他的非阿基米德几何实际上是这样的一种拟几何学,原来以阿基米德命名而现在常常附上欧多克斯(公元前约 350 年)名字的连续性公理不再成立,即两个不同点的横坐标可以只相差一个"真正无穷小的量",它的任何有限倍数都不能等于一个普通的有限实数。

对现代公理理论就简单地谈这些,但在结束之前,我还要讲讲几何公理和定理的本质。当然,这将使我脱离数学的严格范围而进入哲学和认识论的领域。我已经强调过今天绝大多数人在理性上所持有的一致观点,即这里所涉及的是一些主导的概念和公理,必须放在几何学的前面,以便按纯粹的逻辑从中推导出数学证明。但这种一致观点并没有回答关于这些主导概念和公理的真正本质问题。老的

① 第六版,莱比锡和柏林,1922 年。

② 见第一卷《分析》部分,第九章第一节。

观点是，它们是每个人的知觉的所有物，明显而简单，没有人会对它们提出疑问。但这个观点在很大的程度上已被非欧几何的发现而动摇，因为非欧几何已明确表示，空间的直观与逻辑绝不会强行导出欧氏几何（见本节开头部分）。相反，我们看到，在与平行公理相矛盾的假设下，我们得到一个以任意近似程度代表实际知觉关系的、逻辑上封闭的几何体系。不妨说，平行公理是空间关系的最简单的一种表示，它也不过是一种假设。一般来说，基本概念与公理并非知觉的事实，而是对这些事实进行适当选择并加以理想化的结果。例如，一个点的精确概念在我们直接的知觉中是不存在的，它只是想象的极限，是一小块不断收缩的空间在我们心灵上的反映，我们可以逼近它，但又永远不能达到它。

与此相对照，在仅对事物的逻辑感兴趣而对直观或认识论不感兴趣的人们中间，又常常发现另一种观点，认为公理是我们任意作出的说法，而基本概念同样是我们希望用来进行运算的任意符号。当然这种观点有些道理，因为在纯逻辑的内部，没有这种公理与概念的位置，因此必定从其他的源泉，即通过直觉的影响，得到依据或假设。但许多作者过于片面地强调了自己的观点，以致近年来，在现代公理理论方面，我们常常发现自己被误引到了“唯名论”哲学的方向上。唯名论哲学完全丧失了对事物本身以及其性质的兴趣，所讨论的是事物的命名方式及人们使用名称所依据的逻辑。例如说，我们称 3 个坐标的集为一点，并“没有想到任何特定的物体”。而只要某些说法对这些点是成立的，我们就“任意地”提出公理。在这类讨论中，我们可任意地、不加限制地建立公理，只要满足逻辑规律，在公理的完整结构中没有矛盾出现就行了。我不能同意这种观点。我认为这是一切科学的死路。按我的思想方法，几何公理不是任意的，而是可以感觉到的，一般来说，是空间直觉的反映，其精确内容随宜而定。

前面我不断插入一些哲学观点，下面我想补充说明一下几何学的历史，特别是对几何学基础的看法的变迁。与上卷中所说的代数、

算术及分析的历史对比，我们一开始就注意到了几何与它们有很大的区别。上述学科发展到现代的形式，实际上只有几百年的历史。这部历史起自 1500 年左右，即从人们用十进位小数和字母以约数计算开始。但几何学早就是一门独立的学科，其历史可追溯到古希腊时代。它在古希腊时代就已达到很高的发展阶段，以致长期以来，几乎到今天，人们仍把希腊几何奉为一门完整的科学的典范。同时，著名的欧几里得《几何原本》——一部现存的最重要的系统教科书，又一向被看做是希腊数学的全部。确实，在有关科学领域内，很难有别的一本书，能在这样长的时间里保持这样的地位。即使今天，每个数学家还必须使用欧几里得术语。

§21.3 欧几里得的《几何原本》

我首先向你们介绍哥本哈根市 J·L·海伯格整理的欧氏几何版本[1]，从考据学观点来看，这个版本是最好的，其中加上了与希腊文原文对照的拉丁文译文，对于没有学过希腊文的人也很有帮助。确实，欧几里得的希腊文同我们学校里教的希腊文有很大的不同，特别是在术语方面。至于作为欧氏几何入门的文献，我想介绍佐登(Zeuthen)的 *Geschichte der Mathematik im Altertum und Mittelalter*[2] 和马克斯·西蒙的 *Euklid und die 6 planimetrischen Bücher*[3]。如果先读西蒙的书，再读佐登的较一般的讨论，最后读海伯格的版本，你们就会找到入门的途径，不过对海伯格版本应当想尽一切方法仔细阅读，并对译文持批判怀疑的态度。

对于欧几里得的生平，我们知之甚少，只知道他曾于公元前 300

---

① *Euclid's Opera Omnia*，Ⅰ—Ⅴ册，Elementa，莱比锡，1883—1888 年。

② 哥本哈根，1896 年，佐登。

③ 莱比锡，1901 年，也可看 *Abhand lungen zur Geschichte der mathematischen Wissenschaften*，Ⅺ。还可参考 T·L·希思(T. L. Heath)根据海伯格版本译成英文的欧几里得几何原本 13 本及注释，3 卷本，剑桥，1908 年。

年左右生活于亚历山大。但是关于当时亚历山大的一般科学活动，我们是有资料的。在亚历山大帝国建立之后，慢慢产生了把过去几个世纪中所创造的一切加以收集并整理成一个统一的科学体系的需要，因而在亚历山大形成了一种教学体系，相当于我们今天的某几个大学教学阶段。但是当时把手头资料的收集与整理放在优先于使科学研究自由向前发展的地位，因而在整个活动中表现出了学究式的倾向。

在详细讨论《几何原本》之前，让我先对欧几里得或不如干脆说对他的《几何原本》的历史地位或科学意义说几句一般性的评论。尽管要完全了解欧几里得，就要考虑他的为数众多的次要作品，不过我想在这里只讨论他那部伟大的作品也就够了，因为单单那一部作品就已经取得了杰出的支配地位，而从我们的立场来看，这正是迫切需要讨论的。

我要指出，之所以需要做这番评论，是因为对欧几里得《几何原本》的不正确评价的潜在原因是对长期以来广为流传、至今仍深入人心的希腊精神的盲目崇拜。通常认为，希腊文化仅限于相对不多的几个领域内，但在这些领域内精雕细刻所取得的成就，始终是可望而不可即的最高典范。不过现代考据学早已证明这个观点是站不住脚的。考据学告诉我们，希腊人曾孜孜不倦地涉足于一切人类文化领域而不是少数领域，其尽可能表现出的多才多艺，是没有一个民族能相比的。对于他们那个时代来说，他们在每一个领域中的成就确实是值得钦佩的，但是在很多事情上确实也没有超越我们今天认为是最初步的水平。可以说，没有一个领域已达到古今人类成就的最高峰。

特别就数学来说，这种对古希腊的过高估计（或许我应说估计不足），表现在这样一个武断上，说古希腊人对几何学十分注意，已建立了一个不能超越的体系。这种信念特别已造成对欧几里得《几何原本》的盲目崇拜，居然声称这样一种体系已完全实现。与这种陈腐的

信念相反，我愿断言，虽然古希腊人不仅在几何学方面，而且在各个数学领域进行了卓有成效的研究，但我们今天在每个领域，当然也包括几何学领域在内，已经超过了他们。

下面我要详细讨论这个结论，并竭力证明此言非虚。在写《几何原本》时，欧几里得的目的绝不是想编一本集当时几何知识大成的百科全书，否则他不会把当时肯定已了解的一些几何部分完全忽略过去。我只要提到古希腊人已开始进行广泛研究的圆锥曲线和高次曲线理论①就可以了，尽管这种理论到了阿波罗尼斯（Apollonius，公元前 200 年左右）手里才得到充分的发展。而且，《几何原本》只不过是准备对几何学因而也是对数学本身给出一个导论。因此，它看来是为特定的目的而写，只准备讲解柏拉图学派认为必要的数学内容，以便作为一般哲学研究的预备知识。记住了这一点，我们就明白了把重点放在阐明逻辑联系，把几何学当作一个封闭体系来对待，而把一切实际应用抛到一边的原因。为了有利于这个体系，他当然也把当时还没有充分发展、不符合他需要的整整一部分理论知识放过了。

如果与在欧几里得稍后不久生活于公元前 250 年左右叙拉古的希腊最著名数学家阿基米德的个性与成就来对比，我们就可以对欧几里得《几何原本》内容与当时整个希腊数学范围相对而言的局限性得到一个正确的印象。我只提几个特别有趣和突出的事实。

(1) 阿基米德对数值计算表现了强烈的探讨兴趣，这与欧几里得《几何原本》里的主导精神形成了鲜明的对比。事实上，阿基米德的最大功绩之一，就是用逼近圆的正多边形计算了 π。在其他的结果中，他推导出了 π 的近似值为 $\frac{22}{7}$。我只提这一个确定的例子就足以说明了问题。而欧几里得对这样的数值却没有表现出一丝兴趣。相反，我们在欧几里得几何中却只见提到两个圆相互依赖于它们的半径的平方或两个圆周相互依赖于其半径的自身，而对于比例因数

---

① 欧几里得本人也写过圆锥曲线的文章，但没有保留下来。

π,却连计算的尝试都没有。

(2) 阿基米德的特点是对应用的广泛兴趣。众所周知,他发现了水力静力学的原理,并且制造出一些有效的机器,以此积极地参加了叙拉古的保卫战。相反,我们从欧几里得一次也没有提到最简单的绘图仪器——圆规和直尺这个事实,显然可以看出他对应用的关注如此之少。他只是抽象地假设过,通过两点可作一条直线或围绕一点可作一个圆,而对如何去做,却只字未提。在这种地方,欧几里得无疑是处于某些古代哲学学派的统治思想的影响下,以为某门学科的实际应用是等而下之的、鄙夫之所为。遗憾的是,这种观点今天在很多地方仍占上风,始终还有一些大学教师置应用于不屑一顾之地,视之为有伤大雅。对这种傲慢的观点应该予以痛击。每一项值得钦佩的成就,无论是在理论方面或是在实用方面,我们都应当予以同样高的评价。我们应当让每一个人从事他最倾心的工作。这样,任何一个人都会表现得更多才多艺,把他掌握的才能更多地发挥出来。像阿基米德、牛顿、高斯这些最著名的数学家,总是把理论和应用一致包括在研究范围内。

(3) 最后,另一个差别更应引起注意。阿基米德是伟大的研究家和开拓者,他的每一部著作都推动了知识的进步。但是欧几里得《几何原本》所关心的,只是收集手头已经掌握的知识并使之系统化。这就是表述方式之所以不同的原因,去年那一学期我谈到数学的一般结构时曾提请你们注意[①]。在这一方面,1906 年发现的阿基米德的一份手稿[②],是特别具有代表意义的。在这份手稿中,他向一个从事科学研究的朋友透露了他最近对空间图形体积的研究。他的表述非常像我们今天的教学方法。他用演绎法一步步前进,先提示思路,

---

① 见第一卷第一部分附录:《论数学的现代发展及一般结构》。

② 见海伯格及佐登的 *Eine neue Schrift des Archimedes*,莱比锡,1907 年,收于《数学丛书》第 7 卷第 3 册,第 321 页及以后部分。还可参考 T·L·希思整理的阿基米德著作版本,由我译成了德文(柏林,1914 年);那里附有手迹,见第 413 页及以后各页。

绝不使用欧几里得《几何原本》中假设、证明、结论那种僵硬的排列。此外，在这个新发现之前已经了解，古希腊人对系统化课程除了使用"欧几里得"式的这种固定化的表述外，还使用一种自由的演绎方式，当时不仅研究者使用，而且教师在教学上也使用这种方法。或许欧几里得在其他著作以及教学中也使用过这种方法。事实上，那时在亚历山大有过一些类似于我们今天的油印讲义那样的东西，称为"hypomnemata"，即散页的口头表述内容的笔录。

这已足以成为《几何原本》与当时希腊数学全部范围的对比。作为上述讨论的结论，我将通过几个简单的例子来说明，现代数学超越希腊之处已有多远。重要的差别之一是：古希腊人不具有独立的计算或分析学科，既不具有便于数值计算的十进制小数，也不一般地使用计算字母。正如我在去年冬天讲座中所说的，这两者都是近代文艺复兴时期的发明。因此，古希腊人只好用几何形式的运算，用线段或其他几何量的作图运算来代替数字运算，比我们的算术运算麻烦得多。此外，古希腊人也没有负数和虚数，而负数与虚数实际上为我们的算术运算和分析提供了方便的工具。因而古希腊人缺少能把一切可能情况包括在一个公式中的一般化方法。对各种情况的令人望而生厌的细分碎辨，对他们具有重要的作用。这种缺陷在几何学中往往十分明显，而今天借助于分析的工具，正如我们在这些讲座中实际所作的那样，我们能轻而易举地达到高度的一般化。不用区分种种个别情况。指出这几点已足以说明问题，你们可以根据本身的知识，再提出其他许多例子，以说明现代数学超越于希腊数学之处。

在对欧几里得《几何原本》作了这个一般性评论之后，我们现在可以转向具体的讨论。让我先简单地讲一讲《几何原本》第 13 卷即第十三章的内容①。

第 1 卷至第 6 卷讲的是平面几何。前 4 卷包含关于线段、角、面

① 也有第 14 卷与第 15 卷一说(海伯格版本，第 5 卷)；但这两卷非欧氏所作。第 14 卷更可能出自海珀西克斯(Hypsikles)手笔；第 15 卷出自大马斯克斯(Damaskics)之手。

积之类基本几何形状的一般讨论和关于最简单的几何图形（三角形、平行四边形、圆、正多边形等）的理论，讲法和今天相同。与此相联系，第 2 卷给出了几何量的初等算术运算与代数运算，例如两个线段 a，b 之积 $a \cdot b$ 用一个矩形来表示。如果我们今天希望把乘积 $a \cdot b$ 与 $c \cdot d$ 加起来，我们就可以用算术方法立刻算出来，但那时为了再用一个矩形来表示乘积，必须把两个矩形 $a \cdot b$ 和 $c \cdot d$ 变成具有等底的两个矩形。

第 5 卷大大深入了一步，引入了一般正实数的几何等价量，这就是欧几里得称之为“理性的”(λόγos)任意两个线段 a，b 之比 $\frac{a}{b}$。“理性的”这个词是怎么来的，我在上一学期对无理数进行一般讨论时已经提到过（见第一卷第二章）。这个发展的关键是两个比 $\frac{a}{b}$ 与 $\frac{c}{d}$ 相等的定义。这个定义必须具有绝对的一般性，因而当 $\frac{a}{b}$ 是我们所指的无理数时，即当线段 a 和 b 是不对称（欧几里得的说法）时，也应保持。欧几里得的这个所谓不对称，也就是没有公测度，后来就译成“不可通约”。欧几里得的作法如下：他取任两整数 m 和 n，并分别比较 $m \cdot a$，$n \cdot b$ 和 $m \cdot c$，$n \cdot d$ 的长度，从而在 3 个关系

$$m \cdot a \gtreqless n \cdot b \quad \text{或} \quad m \cdot c \gtreqless n \cdot d$$

中，必有一个成立。如对任意值 m，n，在两种情况下总是给出同样的符号，则称 $\frac{a}{b} = \frac{c}{d}$。这完全对应于戴德金引入无理数时所采取的著名的分割过程。

现在欧几里得进一步考虑怎样计算这些比，并发展了他的著名的比例理论，即用以说明型为 $\frac{a}{b} = \frac{c}{d}$ 之方程所有可能的代数变换的几何定理。欧几里得用“相似”这个词来表示比，他的意思是这两对量的“有理性”是相同的。你们看，“相似”这个词的最初意义和现

在的差别有多远！不过在数学中，这个词在某些地方还保持着古义。三角学中现在仍说"纳皮尔相似"，因为必与某些比例有关。说实在的，现在看来没有几个人知道这个名称的真正意义。

比例理论是数学教学中仍然保持着欧几里得传统的一个典型例子。即使今天，许多学校（也许是大多数学校）仍把这个理论当作几何学中专门的一章来教。实质上，我们现代算术已经把它完全包括在内，因此在教这个理论之前已经教了两次——一次是在学习比例时；另一次是在用字母进行计算之初。为什么同一件事要教 3 次，而且要套上特别神秘的几何外衣呢？实在令人费解。这样做对学生必然是很难以理解的。唯一的原因就是人们仍然抱着欧几里得的老的设想不放。但是过去欧几里得所抱的合理目的是想用比例理论来代替他所缺乏的算术手段，对于我们就完全不必要了。

对于今天比例理论处理办法所进行的这个批评，当然并不是贬低欧几里得《几何原本》第 5 卷的科学意义。它的科学意义当然很伟大，因为其中第一次给出了——用现代术语来讲——无理数计算的严格基础，即置于精确的定义基础上。我们在这里清楚地看到，欧几里得的《几何原本》无论过去或现在都绝非像有些人往往错误地认为那样是一本中学教科书。相反，它是针对具有科学思考能力的成熟的读者的。

我必须提到一个传统的观点，认为这个第 5 卷不是欧几里得本人写的，真正作者是克尼多斯的欧多克斯。事实上，人们向来不把《几何原本》看做是一本统一的著作，而看做是把旧有的各个不同部分编在一起的结果。

但不管真相如何，作者究竟是谁，有关的一切资料都如一团迷雾，极难肯定，因为无论是欧几里得或他的任何同代人都绝对没有留下历史注解之类。上述传统的观点出自普罗克拉斯·迪多丘斯(Proclus Diadochus)，他是欧几里得作品的评论者，生活于公元 450 年左右，即比欧几里得晚 700 多年。尽管普罗克洛的论断因种种原

因可能有一定的道理，但我们仍然难以承认它是绝对可靠的，正好像今天有人宣布说，某一部编于 1 200 年前后的作品，其作者为某某人，我们也无法置信一样。

《几何原本》接下去的第 6 卷的内容，是相似形的理论，所用主要工具为比例学说。

在第 7、第 8、第 9 卷中，部分地以几何形式讲解了整数的理论。我们在此发现，对于整数的比即有理数的计算，所用的理论是完全独立于第 5 卷所展开的理论的。尽管有理分数只是实数的特殊的一类，但一点也没有提到更一般的理论。因此很难相信，这两个说法出自同一个作者。在这 3 卷内容中，我只提一下现在用于数论中的两点内容，其一是求两个整数 a 和 b 的最大公约数的欧几里得算法，欧几里得用线段来表示。用现代术语来表示，就是 a 除以 b，然后 b 除以余数，按下列算式 $a = m \cdot b + r_1$，$b = m_1 \cdot r_1 + r_2$，$r_1 = m_2 \cdot r_2 + r_3$，… 继续进行下去，在有限步之后，最终将除尽。最后一个余数就是所找的公约数。另一点是可以在欧几里得这 3 卷中找到存在无穷多个质数的著名的简单证明，我在去年冬天的讲座中已经讲过了(见第一卷第三章:“关于整数的特殊性质”)。

第 10 卷特别冗长，由于几何的表示形式而难以读懂，其中介绍了可以方根表示的无理数的几何分类，是以后在几何构造中要用到的内容。

直到第 11 卷才开始讲测体积学。你们可以看出，欧几里得也不是“融合派”。他尽量把测体积学和测面积学(平面几何)分开，而今天最好像我们常常提到的那样“竭力融合”，尽可能早一点发展整体空间概念，从而使学生一开始就习惯于三维空间图形，而不是人为地把学生初期教学束缚在平面几何的内容上。

第 12 卷中又出现了关于无理量的一般讨论，这对于求棱锥和其他物体的体积是必须了解的。这里，在所谓的“穷举法”中，我们可以看到暗暗地应用了极限的概念，借以严格地推导了无理数之比。这

个方法首先用于证明这样一个求平面面积的定理，即：两个圆正比于其半径之平方。我正好用这个例子来简单解释一下这个方法的内在概念。任何圆可以用不断增加边数的 n 边内接或外切多边形来不断逼近。在多边形面积和圆面积之差会任意小这个意义上，这个圆能被“穷举”。于是，如果得不到比例，则不难理解每个内接多边形小于圆而每个外切多边形又大于圆之间的矛盾(图 21.23)。

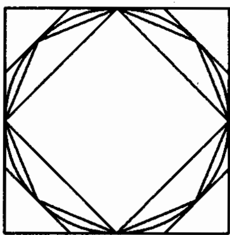


图 21.23

第 13 卷，即最后一卷，包含着正则体的理论，利用第 10 卷中收集的材料，最终证明能用直尺和圆规作出所有正则体即其边长。这个最后的结果与希腊哲学家对正则体所表现出来的一贯兴趣是一致的。

对这些内容作了一般性概述之后，就让我们把注意力集中在欧几里得处理几何基础的那几章。欧几里得的理想目标，显然是从完全预设的一套前提出发对一切几何定理进行逻辑推导。《几何原本》的历史性意义无疑主要在于树立(或传播)了这个理想。但是欧几里得绝没有真正达到他的最高目标。现代科学正是在基本的几何概念上已经取得了更深刻的知识，而在《几何原本》中却发现了一些模糊不清的地方。不过传统的力量是很大的，所以至今(特别在美国)仍奉欧几里得的理论为说明几何基础的卓绝典范。人们把这部著作的历史价值误以为是绝对的和永恒的价值。由于对欧几里得的《几何原本》存在着这种过高估计的情况，所以很自然，在以下的讨论中我要把重点放在它的反面，放在它不再符合我们需要的那些方面。

在进行这种批评中，欧几里得著作原文是否可靠，是一个特殊的困难。大多数原文经普罗克拉斯考证属实，而普罗克拉斯也是最古的考证人。我们所掌握的最古老的手稿都是公元 9 世纪的本子，即比欧几里得晚 1 200 年！此外，这些各种各样的手稿有很大的出入，

而且出入往往正在十分关键的基本部分。拉丁文和阿拉伯文翻译和注释人为了努力澄清原文又出现了许多重大的分歧,形成了一种传统。因此,恢复《几何原本》的真面目,是一个极其复杂的语文考据学的问题,耗费的聪明才智有惊人之多。我们只得迁就这样一个事实:这种考据所得最多只是最接近原文的文字,而不可能是真正的原文。因此,我们从即使最接近原文的许多不同说法中推断出来的东西,绝不可能在一切方面都与实际情况相一致。一般公认,海伯格考证过的文字代表着现代语文考据学的最高水平,我们这些非考据学家最好是以此立论。当然,我们不能忘记,我们所依据的文字也绝不一定与原文一致。因此,如果我们从《几何原本》中找出了什么缺点和矛盾,那么是否应该归之于欧几里得或系误传,都必须始终持怀疑态度。

现在回到主题上来,我们首先要问第一卷中是如何奠定几何基础的。欧几里得在前面提出了 3 组命题,他称之为"定义"、"公设"和"公共概念",用德文来传达,也许可以用 Erklärungen, Forderungen ,Grundsätze 3 个词[1]。对最后一组命题,我们通常按普罗克拉斯的说法用"公理"这个词,不过现在这个词的含义已扩大到包括公设在内。

为了获得关于定义的内容,让我们回忆一下,我们在前面是怎样开始叙述我们的几何进程的。我们说过,我们不能定义点、线、面之类的东西,但我们必须把它们当作每个人都熟悉的基本概念,我们应当加以精确叙述的,仅仅是它们性质中我们希望利用的那种性质。在这种理解下,直到产生解析几何上的坐标系(x, y, z)时,才能进行几何作图。只有在这以后,把 x, y, z 看做是参数 t 的连续函数,我们才能考虑曲线的一般概念。那时我也曾指出,这可能还会包括一般古怪的退化曲线,诸如完全覆盖一个曲面的曲线之类。

---

① 用英文术语来说,就是"定义"、"假设"、"公理"。——英译者注

欧几里得没有这种小心谨慎或从总体设想的精神。他一开始就“定义”了各种几何概念，如点、线、直线、曲面、平面、角、圆等等。第一个“定义”说：“点是没有部分的那种东西”。我们很难承认这是一个真正的定义，因为点绝非仅仅由这个性质决定。接着说：“线是没有宽度的长度。”这种说法甚至是否正确都值得怀疑，如果我们承认前述的曲线的一般概念的话。当然，欧几里得对此是一无所知的。第三个“定义”又说：“一条直线相对于它的点是均匀地排列的线。”这个定义也是完全模糊不清的，可以附之以各种各样的意义。它可能表示该线处处方向相同，这样的话，就必须承认方向是每个人都熟悉的基本概念。也可以解释为：一条直线如果被理解为一个硬杆，在某种空间运动下，即作为轴绕自身旋转或沿自身平移时，它始终与自身一致。欧几里得的这个“定义”实际上以运动的概念为前提。至于欧几里得是否有这个意图，则是一个有争议的问题，以后我们回过来再谈。无论如何，总是无法对欧几里得的直线定义以及其他许多在这里不能详谈的定义找到一个毫不含糊的解释。

我们现在转入对公设的讨论。在海伯格的版本中一共给出了5个公设，其中前3个要求做到：

(a) 从一点到另一点可作一直线。

(b) 有限的直线可无限地延长。

(c) 可以任一点为中心过一给定点作一圆。

我们暂时把第四个公设放到一边而讨论第五个所谓平行公设：

(d) 若两条直线与第三条直线相交且在第三条直线同一侧所构成的两个同侧内角之和小于直角，则两直线向同一侧适当延长后必相交(图 21.24)。

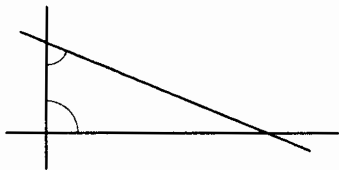


图 21.24

这些公设叙述了某些作图的可能性或欧几里得以后要用到的某些几何图形的存在性。但是还有相当多的类似的几何存在性公

设,虽然也被他用过,却并不能从他已作的公设中逻辑地推导出来。作为一个例子,我要提到两个圆若各通过对方的圆心则必相交这个公设(图21.25)。很容易列出其他许多类似的公设。因此,我们必须说,欧几里得的公设体系是肯定不完整的。

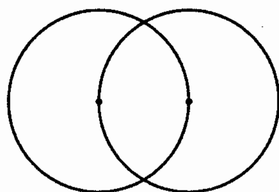


图 21.25

现在让我们来考虑第四个公设。

(e) 所有的直角都相等。

关于这个公设含义何在以及为什么出现在这个地方,一直有很多争论。与此有关的问题是欧几里得是否使用了运动的概念。如果我们前后一致地把刚体图形的运动概念放在开头,像我们在第一种几何发展体系中所做的那样,那么这个公设就是必然的逻辑结果(参考本章第一节),因而是多余的,即使欧几里得并没有这个观点。但在欧几里得的所有这些基本公设中,任何地方都没有明显地提到运动的概念,因此许多注释者假设这个第四公设正是为了引出运动的概念而作,尽管他们都承认这个概念仍不完整。

另一方面,大部分欧几里得评论家认为,欧几里得的本质倾向之一,正是想按照某种哲学考虑(本章第二节开始部分),原则上把运动概念从几何中排除出去,不过那样的话,抽象的全等概念还得放在开头(像我们的第二种发展体系中那样),而这个第四公设就不得不作为全等理论的基础。这就引起一个问题:为什么对于线段的全等没有作类似的说明呢?我们很快可以看出,按照欧几里得体系进一步推导下去,这两个观点都会造成严重的矛盾。

我要指出,欧几里得提出公设时的一般倾向既如上述,则在他的公设中出现这一条公设的原因,就都不能用那两个解释来作出恰当的说明。因此佐登提出了另一个有趣的解释,但也没有充分的说服力。他争辩说,该公设想要说明的是:根据公设(b),过一点可延长一条直线,这个公设是唯一的。佐登的详细解释可在其

所著 *Geschichte der Mathematik im Altertum und Mittelalter* 一书中找到(该书第 123 页及以后各页)。由于始终存在着这个漏洞,因此最后又有一种假设说,这里的原文已被篡改。确实也一再得出过这个结论,无法平息。

现在进而讨论公理。在海伯格的版本中,也有 5 个公理:

(a) 跟同一件东西相等的一些东西,彼此也相等:若 $a=b$, $b=c$,则 $a=c$。

(b) 等量加等量,总量仍相等:若 $a=b$, $c=d$,则 $a+c=b+d$。

(c) 若 $a=b$, $c=d$,则 $a-c=b-d$。

(d) 两个重合的东西是相等的。

(e) 整体大于部分:$a>a-b$。

上述公理中有 4 个在逻辑上是自然的,在此引入的意图,显然是想说明其所表示的一般关系对于一切几何量(线段、角、面积等)也都成立。而第四个公理宣称,关于等或不等的判定标准最终归于全等或重合。这里是否含有运动的概念,也确实不清楚。

关于公理与公设之间的区别,西蒙曾提出一个看法,认为前者处理最简单的逻辑事实,后者处理空间知觉的事实。如果海伯格版本中的顺序肯定与原文相符的话,那么这个看法是十分切合和透彻的。但在各种不同的手稿中,无论是公设和公理的顺序或者是内容,实际上存在着重大的差异,如平行公设往往被排为第十二公理,所以和西蒙的这个设想绝对不一致。

现在我们来更仔细地考察一下建立在这些定义、公设和公理上的欧几里得几何结构的开头部分,即紧跟在公理之后的前四节。通过这四节,同时也可以对欧几里得的基本观点,特别是对运动概念的态度进行某些有趣的观察。

前三节的目的是解决把一个给定的线段 AB 从点 C 开始放到另一线段 CF 上的问题(图 21.26)。当然,任何人都可以用圆规或一

条纸带通过直接转移，即通过平面上刚体的运动来做到这一点。但欧几里得的做法不同，他采用了理论的方法。在他的公设中，他没有假定与圆规的这个自由运动相对应的作图法。根据他的公设(c)(见前述)，只有在圆周上已经给定一个点后才允许围绕一点作一个圆。现在他只能利用公设提供的可能性，因而必须把这个显然简单的作图分解为许多个比较复杂但十分巧妙的步骤：

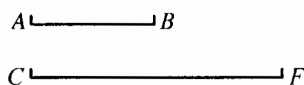


图 21.26

(1) 在给定线段 AB 上作一个等边三角形。公设(c)允许我们以半径 AB 围绕 A 及以半径 BA 围绕 B 各作一圆(图 21.27)。如上所述，可不作任何解释断定这些圆有一个交点 C。由此用适当的公理得出严格的形式的逻辑证明，证明 ABC 是等边的。

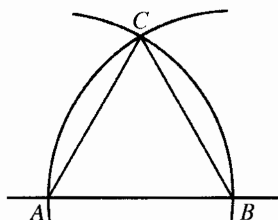


图 21.27

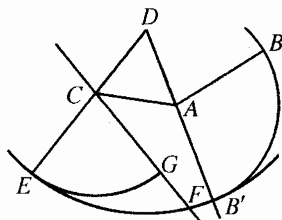


图 21.28

(2) 从一个给定的点 C 作一等于给定线段 AB 的线段(图 21.28)。根据(1)，在 AC 上作一等边三角形 ACD。延长 DA(公设(b))，并与围绕 A 的半径为 AB 的圆(公设(c))相交于 B'(对存在此交点之理由仍未作解释)。现在以 DB' 为半径、以 D 为中心作一个圆，并得出它与 DC 延长线的交点为 E。于是 $CE = AB$，这个显然而见的结果也就得到了详细的证明。

(3) 给出两个线段 AB、CF，满足 $CF > AB$；从点 C 起在 CF 上截一个等于 AB 的线段。根据(2)，从点 C 作任一线段 $CE = AB$，并以 C 为中心、以 CE 为半径，作圆与 CF 交于 G，则 CG 为所求线段。

这就解决了所给出的问题。作为第四个公理，欧几里得现在叙述第一个三角形全等定理如下：如果两个三角形 ABC 和 $A'B'C'$ 各有两边和夹角对应相等（图 21.29，$AB=A'B'$，$AC=A'C'$，$\angle A=\angle A'$），则两个三角形的其他部分均相等。在证明此定理时，根据前述作图，欧几里得犯了一个显然前后不一致的错误，这也是我把全部证明重新写出的原因。他设想把三角形 $A'B'C'$ 叠在三角形 ABC 上面，使得边 $A'B'$ 和 $A'C'$ 分别与边 AB 和 AC 重合，而 $\angle A'$ 则重合于 $\angle A$。现在从所得的结果，我们事实上只学会了怎样把一条线段放在另一条线段之上，但欧几里得只字未提角如何重叠，更未提在这个转移过程中第三边 $B'C'$ 会发生什么情况，连它是否仍为一条直线也存有疑问。直观上，这当然是十分清楚的事，但欧几里得的整个目的是推论的逻辑完整性。尽管如此，他在这里却并未作进一步解释就得出了结论说，$B'C'$ 也必转移成一条直线，因而必与 BC 重合。然而这无异于假设存在不改变几何图形形状与大小的运动，正如第一种几何发展体系中明确所作的那样。如果作了这个假设，当然就能证明这第一个全等三角形定理了（见本章第二节开头部分）。

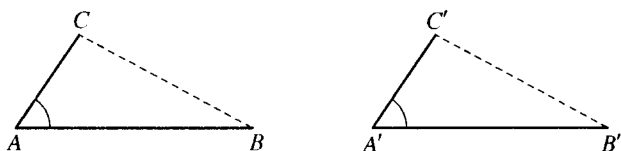


图 21.29

因此，欧几里得的这个证明似乎表明他是支持运动的概念的。问题是：既然如此，为什么在几何基础中只字未提呢？首先，他对第二个和第三个公理的巧妙证明就显得漫无目的，因为这个证明可以利用运动的概念一言以证之。另一方面，如果我们把第四个公理看做是后来插入的，那么欧几里得对第一全等定理持何态度问题仍属悬案。因此在他的体系中仍然存在一个关键的漏洞。没有运动概念是不可能证明这个定理的，因此必须像在我们的第二个发展体系中

所作的那样,把这个定理放入公理之中(见本章第二节开头部分)。作为本讨论的结论,我们只能说,正是在欧几里得几何第一卷前几个定理中表现出来的许多难以解决的重大矛盾,使我们无法说欧几里得几何已经达到了上述的理想。

关于欧几里得对几何基础的表述,还必须提出另一个相反的看法。如果我们用欧几里得自己的理想来衡量他,同时考虑我们现在的知识,那么上述这一切漏洞和模糊不清之处和这一个相反意见相比就没有那么重要了。我们采取熟悉的分析语言,但欧几里得对几何量(线段、角、曲面等)从未使用符号,他把这些量都当作绝对量。他在某种意义上处理了解析几何内容,而其中坐标及其他的量都只以绝对值出现。其结果是他不能得出普遍正确的定理,而必须按具体例子来分别加以说明。提一个简单的例子,所谓推广的毕达哥拉斯定理,其现代表示式为$c^2 = a^2 + b^2 - 2ab\cos\gamma$,一般对锐角或钝角三角形均成立(图 21.30),因为 $\cos\gamma$ 在两种情况下分别取正负值。但欧几里得只知道绝对值$|\cos\gamma|$,因而必须将这两种情况用两个不同的公式

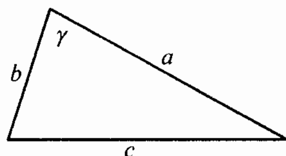


图 21.30

$$c^2 = a^2 + b^2 - 2ab\ |\cos\gamma|\ \text{和}\ c^2 = a^2 + b^2 + 2ab\ |\cos\gamma|$$

来表达。在进一步深入时,这种区别对待的情况当然会变得更为复杂,表达得不够清楚。

我们所谈到的这种缺陷,当然会在纯几何中表现出来。解析表达式中的符号差异,相当于纯几何中一点 C 是在点 A 与点 B 之间还是在线段 AB 之外的顺序差异。只有明确地表达出这个位置关系方面的基本事实,即所谓介于性公理,像我们在第一种以及第二种几何发展体系中所强调的那样,才能建立一个完整的几何的逻辑结构。如果我们像欧几里得那样忽略了这一点,我们就不能达到建立一个纯逻辑几何结构的理想。我们必须继续依赖于图形,并讨论这些位

置关系。因而我们与欧几里得相反的看法，简而言之，在于他没有介于性公理。

必须真正作出某种关于“介于性”概念的假设这个观点，换言之，必须按照某种约定赋予基本几何量以符号的观点，是一种相对来说新的观点。在本卷一开始（第十章中关于多面体体积部分），当我们讨论这个问题时，我说过，1827 年莫比乌斯在《重心计算》中首先先后一致地使用符号规则。在这一方面，1832 年 3 月 6 日高斯致 W · 鲍耶（W. Bolyai）的一封信，十分有趣。不过这封信于 1900 年第一次发表在《高斯著作集》（第 222 页）中，信中说：“为了完全成功，我们必须把‘介于性’之类术语建立在明确的概念基础上，这是完全可行的，但我还没有看到哪里这么做过。”

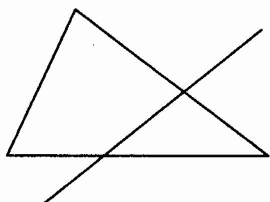


图 21. 31

关于这些“介于性公理”，第一个精确的几何表述是 M · 帕什（M. Pasch）1882 年在 *Vorlesungen über neuere Geometrie*[①] 一书中给出的。这里第一次出现了这样一个典型的定理：若一条直线与一个三角形的一边相交，则也与其他两边之一相交（图 21. 31）。附带说一下，我们在谈第一种几何发展体系中也用过这个定理（见本章第一节）。

这些介于性公理的重要性不可低估。如果我们希望把几何学发展成一门在选择了公理之后不必求助于直觉和图形来推导其结论的真正逻辑性的科学，那么就应该把它们看做和其他公理一样重要。当然，求助于直觉和图形是有启发的，永远是研究中必要的助手。欧几里得由于没有这类介于性公理，始终不得不借助于图形来考虑不同的情况。他把正确的几何作图放到很不重要的地位，所以学习欧几里得几何的学生可能发生因错误作图而得出错误结论的真正危

① M · 帕什，莱比锡，1882 年（1912 年第二版）。

险。大量的所谓几何诡辩就是这样产生的。诡辩定理都有形式上正确的证明，但都基于错误的几何作图，即与介于性公理相矛盾。我提出这样一个诡辩作为例子，你们当中某些人一定也知道，即"证"每个三角形是等腰的。

在中点 D 作角 A 的平分线和边 BC 的垂直平分线。如果两线相平行，则角平分线也就是高，此三角形显然是等腰的。于是我们设它们相交，且按其交点在三角形里面或外面分成两种情况。在每种情况下，都作 OE 与 OF 分别垂直于 AC 与 AB，连接 OB 与 OC。

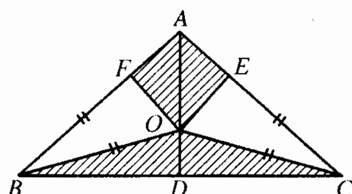


图 21.32

在第一种情况下（图 21.32），画了水平线条的三角形 AOE 与 AOF 全等，因为 AO 是公共边，在 A 处的角相等，且都有直角，因此 $AF=AE$。类似地，带垂直线条的三角形 OCD 与 OBD 是全等的，因为 OD 是公共边，$DB=DC$，对应直角相等，于是 $OB=OC$。由第一个全等性可得 $OE=OF$，从而可推知不带线条的两三角形 OEC 与 OFB 是全等的，因此有 $FB=EC$，将此加于前面方程中，即得 $AB=AC$。

对于 O 位于三角形外的第二种情况（图 21.33），我们用同一种方法指出三对对应的三角形全等，从而求得 $AF=AE$，$FB=EC$，相减后仍得 $AB=AC$，如图所示，从而证明了任何三角形都是等腰的。

这个证明中的唯一错误在于作图，在第一种情况下，O 根本不会在三角形以内，而在第二种情况下，其位置根本不会如图 21.33 所示。两个垂足 E 和 F 中，一个必在其边内而另一则必在其边外，如图 21.34 所示。于是，实际上我们有

$$AB=AF-BF,\ AC=AE+CE=AF+BF,$$

从而绝不能推出两边相等。

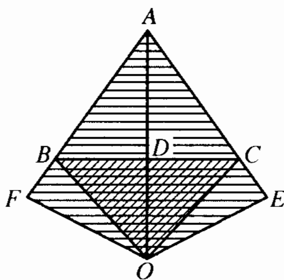


图 21.33

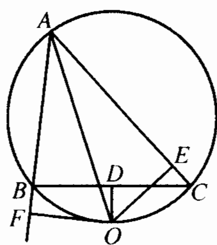


图 21.34

这就完全澄清了这个诡辩。我们也能以同类方法对付其他许多有名的假证明。诡辩的手段总是对错误的图形、歪曲了点和线的顺序。

在批判了欧几里得体系中某些本质缺陷之后，我想指出它的一个最妙的地方，也是他的热心支持者们通常像对他的错误一样视而不见之处。我在前面指出过，在欧几里得《几何原本》第 5 卷中，任何两个几何量 a 与 b 之比，都被认为是一般数量概念的等价物。欧几里得在此明白地规定，只有在某种条件下，即可以找到两个整数 m 与 n，使得 $ma>b$, $a<nb$ 时，才考虑两个同类几何量 a 与 b 之比。他的话是：不同的量只有在其倍数能彼此超过时才有一个比。这个要求现在称为阿基米德公理，但这个名称是与历史完全不符的，因为欧几里得在阿基米德之前就得出了这个公理，而且欧多克斯也许早已了解了。今天，欧多克斯公理这个名称越来越流行了。

这个阿基米德公理作为最重要的连续性公设之一，在现代几何基础研究以及算术基础研究中起着极大的作用，我们在这些讲座中已一再提到它。你们会特别注意到，我们在第一种几何发展体系中所用的"从点 A 出发经过平移而重复的点最终将包括一条射线的每个点"的公设（本章第一节），实质上与阿基米德公理相同。我们在第一卷"无穷小演算的一般考虑"中也详细讨论过这个公理。于是我们称在任意有限倍数 n 之后仍然永远小于 b 的量 a 为相对于 b

的真正无穷小量，或相反，称 b 为相对于 a 的真正无穷大量。因此，欧几里得实际上用他的规定把包含真正无穷小元素或无穷大元素排除在几何量系统以外了。事实上，如果我们希望发展比例理论，就必须排除这类系统，因为正如我们已经强调过的，比例理论只不过是现代无理数理论的另一种形式。因此，欧几里得（或事实上是在他之前的欧多克斯）在这里所做的，本质上就是现代数的概念的研究，而且他使用了完全相同的工具，这正是欧几里得几何中最令人注目的部分。

现在我们来讨论另一个公理，由于这个公理在古代和中世纪时代已众所周知，并已进行过大量的讨论，所以它特别有趣。如果我们考察一下不满足它的一个具体的几何量系统，那么我们就能充分领会它的重要性。我指的是所谓喇叭角，即按某种一般方法考虑的两条曲线之间的角。今天，当我们谈到角时，想到的总是两直线之间的角，而将两曲线之间的角理解为它们切线之间的角（图 21.35）。于是一条曲线，例如一个圆，与其切线之间的角总是零。这样，所有角形成普通的阿基米德几何量系统，我们可以对它应用欧几里得的比率理论，换言之，可以用简单的实数来度量。

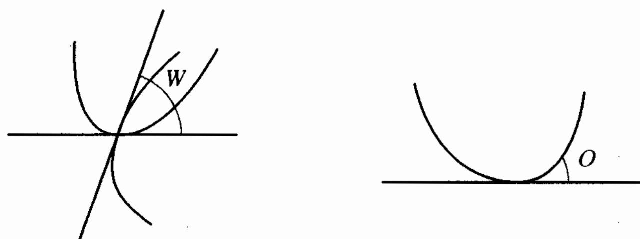


图 21.35

与此相反，我们把两曲线间的喇叭角定义为在交点邻近由曲线所包围的平面的部分（图 21.36）。我们将看到这个定义如何引起非阿基米德量的概念，即不满足那个公理的概念。我们在此只限于角的一边是一条固定直线（x 轴），顶点为原点 O，另一边是与 x 轴相交

或相切于点 O 的圆(在需要时也可以是一条直线)(图 21.37)。于是,当我们向点 O 逼近时,如果一个角的自由边最终位于另一个角的自由边的下边,即前者最终包围之平面部分小于后者,则自然称前者喇叭角小于后者。因此,切圆的角总是小于交圆或斜直线的角。两个切圆中,半径大者的交角较小,因为它在半径小的圆的下面通过。显然,这些规定对任意两个喇叭角决定了谁大谁小,从而所有喇叭角的总体是和我们今天点集理论中所说的与普通实数总体一样是单序的。

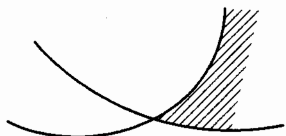


图 21.36

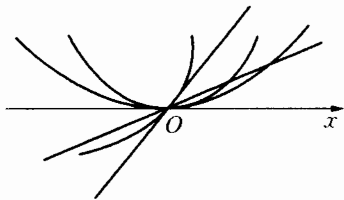


图 21.37

为了判断这两种集合的区别,我们必须对喇叭角的测度作进一步的规定。我们首先用普通角单位来度量过点 O 的直线(与 x 轴)的交角。根据定义,每个由与 x 轴相切的圆所构成的角小于由两直线相交的任意角,但不论多么小,它们都异于零。这种情况在普通数连续统中,对一个异于零的数 a 是不可能的,所以我们的 a 是“真正无穷小量”。

为了把这件事同阿基米德公理联系起来,我们必须对这些曲线角与整数之积给出定义。如果我们有一个半径为 R 的圆切于点 O,则规定半径为 $\frac{R}{n}$ 的切圆构成前者角的 n 倍角应该是自然的。就半径为 R, $\frac{R}{2}$, $\frac{R}{3}$, …的切圆的角越来越大而论,这是符合前面的定义的。因此,一个切圆的角 a 乘以任意整数得出另一个切圆的角,按我们的定义,任何倍数 na,不论 n 怎样大,必须小于一个固定相交直线

所成的角 b(图 21.38),因此,阿基米德公理不被满足,且切圆角相对于相交直线角必须被看做为真正的无穷小。至于两个这种角的一般加法,根据已建立的整数乘法的定义,可将半径倒数相加进行,最终也可以作为真正无穷小角的度量。

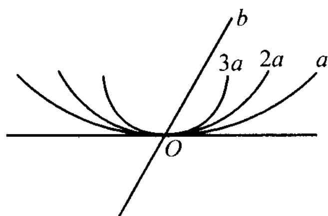


图 21.38

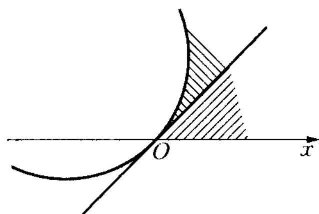


图 21.39

如果我们现在有一个通过点 O 的任意圆(图 21.39),我们可以将它的角当作它的切线与 x 轴的夹角(按通常意义度量)及它与其切线之间的按上面定义的真正无穷小角之和。如果我们对这些分离的被加量施以加法和乘法,我们就对这些喇叭角的运算建立了一个完整的方法。但在这个范围内,阿基米德公理不成立,因而不能应用普通实数。我们可假定欧几里得(及欧多克斯)已知此事,而用他的公理把这种几何量系统有意排除出去。

利用现代方法,我们可以将这些喇叭角的范围推广,使定义变得更广泛而又简单。我们只要考虑所有过 O 的解析曲线,任何这种曲线将由幂级数 $y_1=\alpha_1 x+\beta_1 x^2+\gamma_1 x^3+\cdots$, $y_2=\alpha_2 x+\beta_2 x^2+\gamma_2 x^3+\cdots$ 给出。我们将按 $\alpha_1>\alpha_2$ 或 $\alpha_1<\alpha_2$ 而说曲线 1 与 x 轴的角大于或小于曲线 2 与 x 轴的角。如果 $\alpha_1=\alpha_2$,则相对大小依赖于不等式 $\beta_1\gtrless\beta_2$;如果 $\beta_1=\beta_2$,则用不等式 $\gamma_1\gtrless\gamma_2$ 来决定,等等。显然,用这种方法,可把所有解析曲线的这种角编成一个确定的单序序列,而圆被包含在上面对它们确定的顺序之中。

为了得到曲线 1 与 x 轴的角的 n 倍,我们可简单地令其等于乘幂级数以 n 倍而得到的曲线 $n\cdot y=n\alpha_1 x+n\beta_1 x^2+\cdots$ 与 x 轴的角。

在前面，为了不脱离圆的范围，我们不得不应用一种较复杂的运算，即以半径为$\frac{R}{n}$的圆，其级数展开式为

$$y = n\frac{x^2}{2R} + n^3\frac{x^4}{8R^3} + \cdots$$

代换半径为 R 的圆，其级数展开式为

$$y = \frac{x^2}{2R} + \frac{x^4}{8R^3} + \cdots$$

前一展开式只有第一项是后一展开式的 n 倍。但是采用这个新的、较简单的定义，我们也会得到一个非阿基米德几何量系统。一个展开式从 $x^2(\alpha_2 = 0)$ 开始的曲线，在乘以任意大数 n 倍后，仍然比一个 $\alpha_1 \neq 0$ 的曲线产生的角小。这里我们用一个比较清楚的表达方式，实质上重复了第一卷中的内容①。在幂级数 $y = \alpha x + \beta x^2 + \gamma x^3 + \cdots$ 中，按这里的解释，逐次幂 x，x^2，x^3，…简单地起着不同的、逐级增加阶数的无穷小量的作用。

有趣的是，我们可以通过增加某些非解析曲线更加压缩喇叭角的序列。但为了能对大小作比较，曲线不能无限摆动，或更精确地说，不能和一个解析曲线相交无穷多次。我只要提出曲线 $y = e^{-\frac{1}{x^2}}$ 作为例子就够了。这条曲线具有在 $x = 0$ 处所有导数为 0 的性质，故不能在该处展成幂级数。显然，它最终会在所有解析曲线下通过。尽管我们在前面已有了一个压缩的喇叭角序列，但我们现在有一个新的喇叭角，它和它的有限倍数均小于任意解析曲线和 x 轴所成的角。

我们将在此结束这些讨论及对欧几里得的全部研究。在结束的时候，我要用几句话来总结一下我们经过上述思考分析而得出的对欧几里得《几何原本》的看法。

---

① 见第一卷第九章第一节，其中不同阶的量称为 η，ζ，…。

(1) 欧几里得《几何原本》的伟大历史意义在于把逻辑上前后一致的几何学发展体系的范例传之于后世。

(2) 至于其具体实施情况，有许多做得很好的地方，但也有其他许多地方肯定低于我们今天的科学要求。

(3) 由于原文不肯定，因此许多重要的细节，特别是在第一本开头部分，仍令人怀疑。

(4) 由于欧几里得手头没有算术工具，其整个体系往往显得不必要的赘冗。

(5) 片面强调逻辑过程，使对全书整体及其各本质联系的理解发生了困难。

我想回忆一下我们在不同地方已经指出的两种概念体系，进一步阐明我们对几何发展的态度。

第一是我们可以按照完全不同的体系来发展几何。其中两个体系，我们已经作了仔细的讨论。一个是从运动群，特别是从平移群出发，另一个是从全等公理出发，把平行公理推到很后的地方。这种并列的情况，使我们在确定几何的公理基础时所拥有的自由显得特别重要。我想再次特别强调这个情况，因为经常听到一些不能容忍的说法，作者偏爱某个概念，就把这个概念说成是绝对最简单的、最适合于确定几何的基础，别的概念都不行。事实上，所有基本几何概念与公理的源泉是我们的朴素的几何直觉。我们根据这种直觉选出素材，经过适当的理想化后，把它作为逻辑处理的基础。但对于该做什么选择，并没有绝对的准则。这里所存在的自由只受到一个限制，即要求公理系统满足几何体系前后一贯的需要。

第二是关于我们对于解析几何的态度，以及我们对欧几里得时代以来某些传统的批评。这些传统早就不符合数学科学的现状，因而必须在中学教学中放弃。在欧几里得时代，几何学由于其公理的缘故而成为一般算术（也包括无理数的算术）的严格基础。直到19世纪，算术还保持着从属于几何的地位，但以后发生了变化。今天的

算术，作为一门真正的基础学科，已达到统治地位。在科学几何学的体系中，必须考虑这个事实。也就是说，几何学应从算术研究结果的基础上出发。我们的体系中对解析几何所采取的态度，以及在几何处理中系统地应用分析工具等事实，都表明我们赞成上述主张。

纯几何理论的讨论到此就结束了，希望已经把整个几何领域作了所要求的介绍，没有脱离中学教学的需要。

ementary Mathematics from an Advanced Standpoint

教师应该具备更高的数学观点。理由是，观点越高，事物越显得简单。

《高观点下的初等数学》一书，至今读来仍然感到十分亲切。这是因为，其内容主要是基础数学，其观蕴含着真理，而当时德国数学教育中的不少问题，在今日之我国也仍然存在。克莱因声称本书是为中学教和成熟的大学生写的，但按其内容，所有对数学有一定了解的人都可以从中获得教益和启发……现代数学发生了极大变化，新成果、新概念、新观点、新学科层出不穷。我热切希望我国高水平的数学多面手会写更结合我国实际的、现代化的《高观点下的初等数学》。这样一本书的出版将是我国数学教育史上的一件事。

■ 吴大任

读这本书，您会感到极有收获，而不得不心悦诚服。不得不承认克莱因是真正的大师！

■ 齐民友

除了数学的工作之外，克莱因的数学史至今仍是19世纪数学史上的重要的标准著作，作为当时的领袖数家，他的许多观点至今仍然对数学家、数学史家有所启迪。他的《高观点下的初等数学》反映了他对数学许多观点，是一本译为多种文字的通俗读物，影响至今不衰。

■ 胡作玄

菲利克斯·克莱因教授是德国有名的数学研究家，他也是一位循循善诱的教师。他以罕见的天才，集一数学领域的知识于一身，并善于领悟这一切领域之间的相互关系。他认为使学生了解数学并不是孤立的各学问，而是一个有机的整体，是他作为一个教师的明显职责。他对中学数学教学有浓厚的兴趣，不仅关心该教些什么内容，而且关心怎样教才是最有效的方法……他一贯努力缩短中学和大学之间的差距，从传统漠不关心中激起中学教师对高等数学的兴趣，把中学数学教学引向健康发展的方向；同时也努力扭转大学态度及教学方向，使之承认中学的正常地位，使数学教育前后一贯……《高观点下的初等数学》是一本无珍贵的著作，同样可作为大学教师和中学教师的参考书。无论就材料安排的巧妙或就讨论方式的引人入胜说，目前都没有一本书可以同这本书相比。

■ 洛杉矶加利福尼亚大学数学教授 E·R·赫德里克

■ 伯克利加利福尼亚大学数学教授 C·A·诺布尔

西方数学文化理念传播译丛

丛书主编 汪 宇

Elementary Mathematics from an Advanced Standpoint

高观点下的初等数学

（第三卷） 精确数学与近似数学

〔德〕菲利克斯·克莱因 著

吴大任 陈 鹀 译

 復旦大學出版社

www.fudanpress.com.cn

内 容 提 要

菲利克斯·克莱因是19世纪末20世纪初世界最有影响力的数学学派——哥廷根学派的创始人，他不仅是伟大的数学家，也是现代国际数学教育的奠基人、杰出的数学史家和数学教育家，在数学界享有崇高的声誉和巨大的影响。

本书是克莱因根据自己在哥廷根大学多年为德国中学数学教师及在校学生开设的讲座所撰写的基础数学普及读物。该书反映了他对数学的许多观点，向人们生动地展示了一流大师的遗风，出版后被译成多种文字，是一部数学教育的不朽杰作，影响至今不衰。全书共分3卷。第一卷：算术、代数、 分析；第二卷：几何；第三卷：精确数学与近似数学。

克莱因认为函数为数学的"灵魂"，应该成为中学数学的"基石"，应该把算术、代数和几何方面的内容，通过几何的形式用以函数为中心的观念综合起来；强调要用近代数学的观点来改造传统的中学数学内容，主张加强函数和微积分的教学，改革和充实代数的内容，倡导 "高观点下的初等数学"意识。在克莱因看来，一个数学教师的职责是："应使学生了解数学并不是孤立的各门学问，而是一个有机的整体"； 基础数学的教师应该站在更高的视角（高等数学）来审视、理解初等数学问题，只有观点高了，事物才能显得明了而简单；一个称职的教师应当掌握或了解数学的各种概念、方法及其发展与完善的过程以及数学教育演化的经过。他认为"有关的每一个分支，原则上应看做是数学整体的代表"，"有许多初等数学的现象只有在非初等的理论结构内才能深刻地理解"。

本书对我国从事数学学习和数学教育的广大读者具有较好的启示作用，用本书译者之一，我国数学家、数学教育家吴大任先生的话来说，"所有对数学有一定了解的人都可以从中获得教益和启发"，此书"至今读来仍然感到十分亲切。这是因为，其内容主要是基础数学，其观点蕴含着真理……"。

第三卷目录

第八部分　平面曲线的自由几何

第九部分　用作图和模型表现理想图形

译者的话

克莱因的讲演《高观点下的初等数学》分 3 卷先后出版；前两卷有英译本，经舒湘芹、陈义章、杨钦樑同志译成中文，1989 年出版。第三卷似无英译本，兹据德文本译出。

第三卷的副标题是“精确数学与近似数学”。克莱因认为，数学科学是一个有机整体，包括精确数学和近似数学两部分：“近似数学是数学运用于实际应用的那一部分，精确数学是近似数学赖以建立的坚实框架。”令克莱因不安的是，在他那个年代，“理论数学家的兴趣和思路同数学应用中实际采取的方法离得非常远”。他希望这部分讲演能有助于消除这种分歧，使数学重新成为完整和谐的科学。因此，反复阐明精确数学和近似数学的相互作用及其区别，批判在这个问题上（特别是教材里的）种种错误观点和错误做法，就成为贯穿全卷的中心思想。

克莱因提醒人们，不要因为近似数学这个名称而降低它的地位。他指出：近似数学不是“近似的数学”而是“关于近似关系的精确数学”。他又指出，“必须从精确数学中区分出那些对实用没有直接意义的东西”，诸如无理数概念，有公度与无公度的区别等等。至于所谓的“几何三大问题”，当然都完全属于精确数学。

本卷讨论的内容侧重几何方面，采用的是解析方法。联系着本卷的中心思想，克莱因着重论述了精确几何与应用几何、理想图形与经验图形、几何理论和直觉的关系。他指出：精确几何的基础是公理系统，采用解析方法时，相当于现代实数概念；但公理产生于感觉（和人的创造），而且直觉还引导人们逻辑思维，它又是精确几何新发现的源泉。他同时强调，严格的证明不能单纯凭直觉。

和前两卷类似，本卷着重讨论一般教材不讨论的内容或是它们所忽略的精微处。例如在精确数学方面，克莱因详细介绍了魏尔斯特拉斯连续而无处可微的函数，以及作为闭节的单值连续像而完全覆盖一个区域的皮亚诺曲线，讨论了二元函数的连续性和偏导次序可颠倒的条件。在近似数学方面，他突出了函数带概念，详细地介绍了插值法，并随时把误差估计放在十分重要的位置。他通过一些植根于实际应用的反演群来获得某些具有值得注意的性质的点集，如一个无处稠密却自稠密而完备的点集，一条处处有切线但非解析的若当(Jordan)曲线。他用一定篇幅论述了应用几何中的测量学、作图法和几何模型。克莱因认为，“在测量学中，近似数学的思路发挥得最清楚而彻底”。结合作图法和几何模型，他论述了当时代数几何中的一些基本概念和若干引人注目的成果。

本卷涉及的数学内容一部分是较复杂的，但由于采用由简到繁、由具体到一般的论证方法，而且到处结合几何图像，读起来仍然感到引人入胜。和前两卷一样，克莱因成功地把严谨的科学态度和避免过多形式推证、灵活多彩的阐述方法巧妙地结合起来。

克莱因不愧为“塑造几何图像的造型艺术大师”，他十分热心地鼓励图片和模型的制作，而结合图片和模型来讲授与学习几何和一般数学，其效益显然是巨大的。制作图片和模型需要学识、想象力、技巧以及不畏繁琐的奉献精神。电子计算机对这项工作提供了极大的便利和新的无限广阔的可能性。例如制作关于数学的动画片，将比以前容易得多。这项工作在中国还很薄弱，我们希望有志之士会为此努力。

原书底注有一部分带方括号的是第三版整理者赛法特加的，中译本又加了一些中译者注。希望读者对中译本中的错误，无论是内容还是文字方面的，都不吝指出，以凭订正。本卷中译本在付刊期间承陆秀丽教授鼎力参与校订，十分认真仔细，减少了不少差错，谨此致谢。

译　者
1991 年 7 月 1 日

第 一 版 序

去年夏天我所作的讲演已经以手稿形式印出来了，现在我愿意沿着我近年所主要致力的方向迈出新的一步：把数学科学重新结合成为包罗它的各个方面的有机整体，来发挥其作用。早些时候，在专业分支还没有出现时，数学自然就是这样的；现在不同了，特别是像目前那样，抽象数学和应用数学的代表人物越来越明显地站在互相对立的地位。针对最后这一种情况，我已经反复指出，必须对这两部分数学的区别及其相互联系作出明晰的阐述，并把这作为有核心意义的一项工作。以后数学的这两部分将称为精确数学与近似数学。布克哈特和霍伊恩(Heun)两位先生近年都发表过和我类似的见解，可以参阅 Burkhardt 先生的就职演说[苏黎世，1897 年①]，它在 *Jahresberichts der deutschen Mathematiker-Vereinigung* 第 11 卷第一册重印，从而使更大范围的同行容易看到。还可以参阅霍伊恩先生关于科技中的动力学问题的报告，它发表在上述杂志第 9 卷(1900—1901 年)。在这篇报告中，就我所知，第一次出现“近似数学”这个词，至少我是从那里取用这个词的。我希望看到对上面所说的互相对立的现状有详尽的分析，像下面的讲演在几何领域中的作法，以引起更多人士对这个问题的兴趣和理解。

在这个讲演里，我同时对以前关于数学教育，特别是高等学校的数学教育所作的各次讲演(参看，例如 *Jahresbericht* 第 8 卷，1898—1899 年)加以充实。以后和以前一样，我的观点是，对初学者以及对

① 数学和科学思想。

那些把数学仅仅作为进一步学习工具的青年学生的教育，从感性阶段起，就要让他们去进行简单的应用。从教学观点来看，这样做是必要的，这是近年从对大多数学生的调查研究中越来越明显地看出来的。在外国也是如此。但是，并非次要地，我有一种信念（而且从不放弃它），认为在当前数学发展的形势下，对于培养高级数学人才，这样的教育还不够，在直观事实之外，还必须强调现代实数概念的核心地位以及和这个概念联系着的深入发展。遗憾的是，就我所知，在教科书和讲义中，还缺乏从这一观点引导到那一观点的论述。这个讲演将可满足这方面的部分要求，它的最高目标是，有朝一日它要成为多余的。因为那就表示，它所阐明的意见已经成为高等数学教育中理所当然的组成部分。

F • 克莱因

哥廷根，1902 年 2 月 28 日

第 二 版 序

第二版基本上是这个讲演 1902 年第一版的重印本，改动极小。只是改善了个别不准确处，并补充了一些新发表的文献，这些文献都是涉及与讲演内容密切相关的发展的。最后还重印了“Gutachtens der Göttinger philosophischen Fakultät betreffend die Beneke-Preisaufgabe für 1901”（“哥廷根哲学系关于 1901 年本内克奖课题的报告文集”），本演讲内容有多处和它有关。

米勒（C. H. Müller）

哥廷根，1907 年 1 月 5 日

第三版序

作为F·克莱因的《高观点下的初等数学》第三卷，这部讲演在“精确数学与近似数学”(Präzisionsund Approximations-mathematik)的副标题下初次以书本的形式出现。这是在前两卷出版几年前克莱因所作的讲演。由于其目标和内容，它也像前两卷一样，拥有广大读者。作为手稿，它长期被引用的标题是 *Anwendung der Differential-und Integralrechnung auf Geometrie* (*Eine Revision der Prinzipien*)(《微积分对几何的应用(对原则的一项修改)》)。改变标题是F·克莱因本人的意愿，在他去世前的两个月里，我和他关于出版这一卷的需要曾有过一系列的交谈。克莱因认为，新标题比原先标题更合乎讲演的倾向。

编辑这部讲演所采取的基本方针和出版前两卷相同。原来的论述总的来说都保持不变，但在许多具体地方由于内容和形式的需要，做了改动和补充。一部分补充当然是考虑到讲演本身的需要而添加了一些新文献，其他补充则以底注形式出现(放在方括号内)，插图改善很多，特别是在讲演的第三编里，要增加一些图来刻画空间的三次曲线和三次曲面的形状。在那部分讲演里，显示了克莱因是塑造几何形象的造型艺术大师，而在这方面他是很闻名的。原来写这部分时，他设想读者手边有相关的模型。通过增加插图和补充内容读者可以不必依赖模型。在充实这部分时，克莱因 1907 年所作关于曲面和空间曲线的讲演起了很大作用。

1907 年第二版中加进来的“哥廷根哲学系关于 1901 年本内克奖课题的报告文集”现在删去了，因它已重印在克莱因的数学论文集

的第二卷里。

在编辑工作中，我得到哥廷根的瓦尔特先生许多有价值的建议和帮助。他和科隆的弗尔梅特（H. Vermeit）先生都读了校样，后者还负责编制了索引。我的同事霍曼（H. Homann）协助我拍照了当地收集的一系列数学模型。特别是柯朗（Courant）教授先生和我进行了多次有关我工作进展的谈话。对于上述先生们的帮助，我应当对施普林格出版社热情地满足我的愿望表示深深的谢意。

赛法特

哥廷根，1928 年 1 月 31 日

前　　言

从最近文献中可以看到一种深刻的分歧：理论数学家们的兴趣和思路同数学应用中实际采用的方法，离得非常远。对此，你们必定感到惶惑。这不但损害对数学人才的培养，也损害数学科学本身。看来，反对这种弊端是非常重要的。我现在开始的讲演，希望能对此做出贡献；我试图从所谓认识论的观点，把各种不同数学问题彼此之间的关系，按其本来面目加以阐明。你们既要学会了解现代理论家们的兴趣，又要能判断数学思维中哪些部分对应用是有直接意义的。我不怀疑，你们对于已经出现的不同观点的对立有兴趣并由此获益。希望我的讲演导致这样的结果：你们将来会做出贡献，以改变数学科学目前存在的不良的发展片面性，使数学重新形成包罗各个方面的一门和谐的科学。

我提出的计划太广泛，在一个学期内不可能对一切方面都加以论述。因此，我将主要突出数学中一个领域，即几何学。几何的实践包括几何作图和度量，这些问题，从希腊人开始，都已用抽象的处理方式在理论上解决了，我将采用分析方法论述从几何作图和度量两方面提出的问题。这样做本来不是必要的；人们也可以完全不从几何领域出发。不过，我所提出的问题基本上是结合分析学的；特别是联系到微分和积分的发展及其作用的。

关于文献，我不能举出教材。我将在有关地方给出单个的文献；此外，我愿意推荐 *Enzyklopädie der mathematischen Wissenschaften*（《数学科学百科全书》）。这部书的目的，在于全面论述19世纪数学，并对全部有关文献加以编列。

若要把同样的计划推广到数学的其他领域，就特别要考虑数学在自然界研究中的应用(如力学等)。与此有关的一系列想法已发表(在报告文集里)，这是 1901 年我们哥廷根哲学系在纪念本内克(Beneke)基金(Beneke-Stiftung)时的报告集[①]。此外，主要可以举出 1873 年我已发表于 *Erlanger* 报告中的一篇文章，题目是"Über den allgemeinen Funktionsbegriff und dessen Darstellung durch eine willkürliche Kurve"("关于一般函数概念及其用任意曲线的表示法")[②]。还有我在 Evanston 讨论会上所作的讲演"Lectures on Mathematics"("关于数学的讲演"，纽约，1894 年[③]。由兰格尔(L. Langel)译成法文，巴黎，1894 年)中的第 7 篇报告，题目是"On the mathematical character of space-in-tuition and the relation of pure mathematics to the applied sciences"("关于空间直觉的数学本质以及纯粹数学对应用数学的关系")。

---

① *Math. Annalen* 第 55 卷(1902 年)，翻印在 F. Klein: *Gesammelte Math. Abhandlungen Bd.* Ⅱ(《数学论文集》Ⅱ，第 241—246 页)。

② 重印于 *Math. Annalen* 第 22 卷(1883 年)(以及 F・克莱因:《数学论文集Ⅱ》，第 214—224 页)。

③ 美国数学会新版，纽约，1911 年(所述报告又见 F・克莱因:《数学论文集Ⅱ》，第 225—231 页)。

第七部分

实变函数及其在直角坐标下的表示法

第二十二章
关于单个自变数 x 的阐释

§22.1 经验准确度与抽象准确度，现代实数概念

我将以下述方式系统地开始：先完全不提函数，而试着重点考虑自变数 x，并像通常那样，用一条坐标轴上的点代表它，这个点和原点的距离用长度 $|x|$ 表示，而且当 x 有正值时，点在原点右边，当 x 有负值时，点在原点的左边。我想集中注意力于作出这样一个点的准确度。这样，我发现，当我采用任意一种方法来作图或度量，通过肉眼的视觉作用，甚至把空间直观的抽象也考虑进去，要用某个数来代表一个点，其结果的准确度总是有限的。

利用目前最精密的显微镜和显微度量仪，测量 1 米长度的精确度实际上不能超过 0.1 微米。若想再进一步，就会遇到光的衍射作用对最精密的显微镜的制约。因此，关于光和物质的相互作用的物理定律，就阻止了度量的精确度进一步提高①。用米表示，0.1 微米 $=10^{-7}$ 米，所以，用米表示，长度的直接测量，充其量只能达到小数点后第 7 位。

用肉眼作图和度量的情况也与此类似，不过这时临界值，即观察的精确度所不能越过的限度，自然要大得多。所以，可以叙述一个最

① 这是作者当时的情况。随着科学技术的发展，现在测量的精确度已大大提高。但作者的观点本质上仍然正确；因为任何实际的测量，其精确度总要受到所采用手段的制约。——中译者

高定则(不是数学定理,但可以作为数学应用中的基本定则)如下:在每一个实用领域中,必有精确度的一个阈值。

现在,这同空间感觉是怎样相联系的呢?

我们现在来到一个有争论的领域中;因为不同时代的哲学家们对此提出了最为分歧的观点。特殊地,许多哲学家把空间感觉看成是绝对精确的。与此相对立,需要强调的是,现代数学家却能够举出许许多多不同的空间构造,而由于它们所涉及的结构异常细微,要从这些空间构造形成空间感觉是根本不可能的。因此,我相信,对于空间感觉也有一个阈值。但是,如果你们不接受我这个观点,你们可以从不同角度来考察那些空间构造,也尽可以亲自尝试看能否从它们建立空间感觉。关于空间感觉,我只提出以下定则:尽管对于空间感觉是否存在阈值的问题还可以争论,对于后面所出现的例子仍然可以反复考察验证。

现在,如果把确定 x 的数值的经验方法和实数理论中关于数的算术定义相比较,就会发现,根据算术定义,x 的值的精确度是没有限制的。

让我们先来理解一下实数的算术定义,简言之,关于现代的实数概念①。

为了避免过分详细的陈述,可以说,我们把全部十进制数,看成全部实数的代表。在这里,十进制无尽小数也作为一定的数的代表。可是相反,是否每一个实数也能用唯一确定的十进制数代表呢?

在这里,我们指出,有以下定理②:每一个有尽十进制数,可以换个样子写,即把它最后的一位数字减 1,然后在它后面加上一连串的 9。不过,可以证明,这是把实数用十进制数代表时所遇到的唯一的不确定性。所以,每一个有尽或无尽十进制数对应于一个实数;而每一个实数,除了上述例外,都只对应于一个十进制数。

---

① Modernen Zahlbegriff 直译是"现代数的概念"。——中译者

② 参考第一卷第 34—37 页(中译本第 24—26 页)。

在这个表示法中,如何区别有理数和无理数? 从初等数学,我们知道答案是: 一个有理数 $\frac{m}{n}$,其中 m 和 $n\neq0$ 是整数,是一个十进制数,它或是中断的,或是周期的[①]。附带一提,这个结论的逆也是正确的。

当然要假定,通常用于有理数的运算规则可以推广,也适用于无理数。我们特别要推荐 K · 克诺普(Knopp)的 *Theorie und Anwendung der unendlichen Reihen*(《无尽级数理论和应用》,第二版,柏林,1924 年),那里有明晰而深入的关于无理数的论述。此外,可以参考《数学科学百科全书》第一册中 A · 普林斯海姆的一篇报告"Irrationalzahlen und Konvergenz unendlicher Prozesse"("无理数与无尽过程的收敛性")第 49—146 页,首先是第 49—59 页[②],每一个数学家都应该阅读它。

§ 22.2　精确数学与近似数学,纯粹几何中亦有此分野

现在回到我所要特别强调的一点: 用十进制数来代表实数是绝对准确的。十进制数决定实数本身,所以在抽象算术里,上面所说的阈值就无限制地缩小。这样,我们就有关于经验与理想化之间的基本的对立定则。对于当前这个特殊情况,我把它叙述为:

在算术这个理想领域里不存在像经验领域里那样的不等于零的阈值,在确定实数(或者作为确定实数的定义)中,其精确度是没有限制的。

这里所说的在实用几何里通过经验与在抽象算术里通过精确的定义来确定一个量,其区别在于有限制的和无限制的精确度只是一

---

① 即有尽位数的或循环的。——中译者

② 这篇报告 1898 年结束。还可以参阅 A · 普林斯海姆的 *Vorlesungen über Zahlenlehre*(《数的理论讲义》),莱比锡。

种特款。若把外界感知或者应用操作的任何领域和抽象数学相比较，这类性质的区别就总要反复出现。对于时间，对于一切机械的和物理的量，尤其是对于数值计算，都是如此。用七位对数表来运算同只准确到七位数的近似计算其间有什么区别？另一方面，我们将看到，在每个领域里，通过适当的公理，人们可以达到绝对精确性；于是我们恰恰是在经验事物上建立理想思维的事物。

绝对准确性与有限度准确性之间的这个区别，将要作为一条主线贯穿于整个讲演。可以通过它把整个数学一分为二。我们区别：

(1) 精确数学(只用实数运算)；

(2) 近似数学(用近似值运算)。

近似数学这个词并不意味着要降低这个数学分支的地位，因为它甚至也不是一种近似的数学，而是关于近似关系的精确数学。只有把两部分都包括在内，才有完整的数学科学：

近似数学是数学科学运用于实际应用的那一部分，精确数学可以说是近似数学赖以建立的坚实框架。

我首先通过数值运算来说明数学的一分为二。在下面这个例子中，上述的矛盾是直接看出来的，而不是被复杂的心理作用冲淡了的。

在精确数学里，我们仅仅运用实数，如上面所说，它们是用有尽或无尽十进制数代表的。那么，例如，实数 $x=6.437\ 528\ 4\cdots$ 中，已经准确到第七位小数，如何表示这个近似值呢？早在18世纪，数学里已经用函数记号 E 作为法文的 entier 的缩写。$E(x)$ 是含在 x 中最大的整数[①]。利用这个记号，上述的近似值显然就可以写成 $\dfrac{E(10^7x)}{10^7}$，用它代替 x 来运算[②]。以精确数学的概念结构作为准绳，就可以说，在近似数学的数值运算中，所处理的根本不是实数 x，而

① 即在 x 的十进制数表示中的整数部分。——中译者

② 若不简单地截去那个十进位数的尾巴而按四舍五入的方法处理，则 $E(10^7x)$ 可以看成是在 $10^7x-\frac{1}{2}$ 和 $10^7x+\frac{1}{2}$ 之间的最大整数，其他一切不变。

是与实数 x 有关的 E 函数。

与此完全类似，经验几何和抽象几何之间的区别也可以改写如下：

首先读读实用几何的性质。它是当我们要处理具体空间关系时(例如作图和造型以及度量中)所用到的几何，在这里，一切操作的准确度都要受到某些阀值的限制。因此，在应用几何这个学科中，可以提出下面的定义和指导原则：

在这个学科里，一个点是一个实体，它的范围小得可以忽略。

一条曲线，特殊地一条直线，是一条带，它的宽比起它的长实际上是微不足道的。

两个点决定一条连接它们的直线，两点距离越远，连线就确定得越准确。若它们紧紧靠近，这条直线就很难确定。

两条直线决定一个交点，它们的夹角越接近直角，交点就确定得越准确；夹角越小，交点就越不准确。

现在谈抽象几何：在这里，没有上述那样的不确定的命题。这是因为，抽象几何，是从公理出发的，这些公理以绝对的形式出现，而对于应用几何，这些公理只是近似地适用的。因此，这里有着截然不同的思路。在实用中仅仅近似地正确的关系成为严格地正确的假设，并且在这些“公理”的基础上，通过纯逻辑的推理来获得抽象几何中的结论。于是，和上列的几条相应，这里规定：

一个点没有空间的延伸。

一条线有长无宽。

两点有一条完全确定的连线。

两条相交直线总有一个确定的交点。

与此相关，我特别愿意提请你们注意的，是这两种完全不同的几何之间的联系(相对于刚才在算术中引进 E 函数的作法)。这种联系的存在，是每一个人都默认的。尽管抽象几何的产生，以实用几何为基础，抽象几何的结论却总要转过来应用到实际上，这只是从理性认识上说的，但还必须考虑一种情况，即观察误差理论。每一项观察

或实际操作都会有误差,更正确地说,按照抽象几何观点,每项观察不是给出一个确定的量,而只是给出在一个幅度里的量。若问抽象几何以什么形式运用于应用几何,答案是:

当人们已经知道一些数据是在某个幅度中变化的时候,就要研究以这些数据为基础的数学推理结果,要在怎样一个幅度里变化。

我现在回到我们整个讨论的始点,在那里,我们令实数和一条坐标轴上的点相对应,并且指出,在实际作出这个对应步骤时,只能达到有限度的准确性。若要有更准确的规定,就要有不受经验支配的公理。而这样就要进入抽象几何的领域。

我们提出以下的公设:每一个(在上面的意义下的)实数 x 对应于坐标轴上唯一的一点,而每一点对应于唯一的实数。

这个公理对于解析几何来说是基本的;若要把它和理论精确性联系起来,也可以简单地说,它对于精确几何[①]是基本的。我并不否认直觉推动我们去提出那样的规定,但要指出,我们的公设是从有意义的感觉和我们的创造力产生的。我这里的意思可以通过两个例子来说明,而这两个例子是从集合论选出来的,人们公认它们对当代数学起着富有意义的作用。

按照我们的定义,一个数集也是在坐标轴上的一个点集。它是这样一些数或点 x 的整体:对于每一个数或每一个点,总可以判断它是否属于那个整体。

试举两个简单点集的例:

(a) 一切从 0 到 1 的有理数 x 或者点 $\frac{m}{n}$。

这个点集在线段 0…1[②] 上"处处稠密"。线段上任意小的子节[③]里总含有无穷多个有理数,粗糙地说,这个子节就像一个布满漏洞的筛

① [上述公理称为关于直线连续性的康托-戴德金公理。]

② 本书中用 0…1 表示以 0 和 1 为端点的节。——中译者

③ "节"有人称为"区间","子节"称为"子区间"。——中译者

子,因为里面总还有不满足我们点集定义的无穷多个无理数。这里就有这样一个问题:根据我们的空间直观,能够设想一个处处稠密的点集还不能填满那个连续线段吗?但按照抽象定义,这却是清楚的,我们那个处处稠密的点集就是不能填满那个线段。

(b) 一切从 0 到 1 的实数,但不包括上限 1 和下限 0,即一条去掉端点的线段。

按照我们的空间直观,似乎不可能有那么一条没有端点的线段。所以,不包括上、下限的一切从 0 到 1 数的整体,不能在空间实现,可是它却是完全确定的一个集合。

从这些观察,我们得到一个重要方法论上的结论如下:

由于抽象几何的事物不能完全在空间的直观实现,在抽象几何里,一个严格的证明绝不能纯粹以直观为依据,而必须从已经假定为绝对正确的公理出发来进行逻辑推理①。尽管如此,另一方面,直观甚至对于精确几何仍有着逻辑推理所不能代替的价值。它引导我们进行论证,此外,它又是新发现和联想的源泉,下面的就职讲演可作参考:

Hölder, O.: *Anschauung und Denken in der Geometrie*(霍尔德:《几何中的直观和思维》),莱比锡,1900 年。

§ 22.3　直观与思维,从几何的不同方面说明

上面我阐明精确数学和近似数学的区别,并且强调,它们在一起才构成整个数学科学。当前数学科学工作中出现的缺憾是,理论家们过分片面地对待应用数学,同时实用家使用近似数学时,又不以精确数学为指引,不求助于精确数学来推动近似数学前进。

---

① 大概现在每个理论数学家都会同意这个提法,可是我觉得,只有像我所讲的那样,阐明空间直观不准确性,才能使人信服。

关于这里所说的理论家和实用家的对立，我还要用另一个例子来说明，它取自通常的初等几何教学。

从欧几里得开始，绝大多数人都认为初等几何是精确数学，它从某些公理(这些是作为完全正确的公设提出来的)作为最高的准则，然后从它们通过纯粹的逻辑方法推出新的定理和证明。因此，必须从中区分出来那些对应用几何没有意义的东西。

在这方面，我首先指出有公度和无公度线段的区别。请注意，当我们的观察只具有有限度的准确性时，这种区别是根本没有意义的。例如边长为 1 的正方形的对角线是 $\sqrt{2}=1.414\cdots$对于这个无尽小数，可以用 $\frac{14}{10}$，$\frac{141}{10^2}$，$\frac{1\,414}{10^3}$ 逐步逼近，也就是可以用两个整数的商来逼近使它达到任意小的已给阈值。比如，若令阈值等于 $\frac{1}{10^6}$，就可以满足度量几何中的每一个实际需要，超过这个限度，就只有理论家才感兴趣。

有公度和无公度之间在严格的形式下的区别纯粹属于精确数学(无理数的概念也是如此)。

但是，当实用家不把“有公度”和“无公度”这两个词从他的术语中完全排除在外的时候，他们是在什么意义下使用这两个词的呢？我目前主要是想到天文，特别是其中的摄动理论。试取木星对一个小行星，例如 Pallas 的影响作为例子。当天文学家说，Pallas 和木星的周期有公度，是什么意思？无论一个行星的周期有多大，由于存在着不同类的摄动影响，它的确定根本上只能有有限度的准确度，即对于通用的单位，它只能准确到一定位数的小数。所以，当天文学家说，两个行星的周期是有公度的，他的意思只是说：所得到的周期的比值可以用两个小的整数之比来表示。对于木星和 Pallas，可以表示为18∶7。但是，如果有一次观测或计算所得到的比值是，比方说，18 000∶73 271，人们就不能再说，这两个周期是有公度的了。

现在回到欧几里得几何学，我们提出这样的问题：一个几何问题能否用圆规和直尺来解决。即能否通过有尽多次运用这两种工具

来得到一个解。对这个问题,我们必须回答:

一条线段能否用圆规和直尺作出的问题,也是属于精确数学的范畴。它关系到在公设的假定下能否有一个精确解的问题。

在这方面,我们知道从古代起,就留下了 3 个据猜测都无法解决的重大问题,它们的实际不可解性是近代才证明的。它们是:

(a) 任意角的三等分。

(b) 立方倍积,即由已给单位长线段作 $\sqrt[3]{2}$。

(c) 化圆为方,即由单位长线段作 $\pi=3.1415926\cdots$

上面我已经指出:(a),(b),(c)3 个问题之间不能通过有尽多次用圆规和直尺来解决,是属于精确数学而不属于近似数学的范畴。在实用上,我们可以把它们中的每一个用圆规和直尺解决到实际可能的准确度①。例如很容易用外切多边形和内接多边形去逼近一个圆;从而逼近 π;但也存在着远为简单而且往往是非常有意义的逼近法。每每有一种纯粹实用的步骤。为了把一个角分成三等分,我们宁可试下去,一直到分得充分精确,这个方法比下面的方法实际上要好得多。那种方法就是先计算到一定的小数位数,然后把解答搬到图上(可以参考地质学和天文学上用来划分圆的仪器②)。

§ 22.4　用关于点集的两个定理来阐明

现在再回到关于点集的讨论。我们已经得到这样的结论:在精确数学的证明里,不能简单地依靠直观,尽管图形在宏观上对于有关证明的启发是非常重要的。下面我给出两个关于点集的非常简单的

---

① [关于这 3 个问题的不可解性,可参考本书第一卷第 56—60,123—124,256—269 页(中译本第 49—53,125—126,274—288 页)。关于逼近方法,可看内容特别丰富的一本书,Th. Vahlen: *Konstruktionen und Approximationen*(范伦:《作图与近似法》),莱比锡,1911 年。]

② [在 L. Ambronn: *Handbuch der astronomischen Instrumentenkunde*(阿姆布朗:《天文仪器知识手册》)第 1 卷(柏林,1899 年)第九章,第 433—437 页。]

定理，上面所说的就适用于这两个定理。可以看出，实际上引出它们的证明是什么。

定理 1　假定在坐标轴上，已经给定一个无穷点集，它不超越一个已给点 A（它不需要达到 A），这样的点集称为有右（上）界的。要证明的是，这个点集有一个最小右（上）界点。

定理 2　在坐标轴上，假设已经给定一个在 A,B 两点之间的无穷点集，则在 A,B 之间，至少有一个聚点，即在这个点的任意小的邻域里，总有点集的无穷多个点。

当人们第一次听到这些定理的时候，它们也许显得是自明的。它们的确是自明的，不过是在较高的意义下，即在数学认识整体的观点下自明；因为从根本上来说，数学不过是自明事物的科学。对于这两个定理，我们不会也不能提供令人惊异的东西，我们只是会清楚地看到，这些定理如何联系着在我们的论述中居于首要地位的实数概念。

如何看待我们的第一个定理?

定理中那个点集是完全确定了的，即一个点是否属于点集是确定的。现在，点集里可能有一个点确实有最大的坐标，这时它显然就代表所求的最小上界。这样的点集的例有：

(a) 一切 0 与 1 之间的点，包括 1 在内；

(b) 上述点集的一切点，再加上点 2。

在前一款，点 1 是最小上界；在后一款，点 2 是最小上界。但是点集可能是这样的：它不含有位于最右的一点，即当我们设想把它的点按坐标大小排列的时候，尽管它们不越过坐标轴上的某一点 A，但点集里却没有具有最大坐标的点。这样点集的一例，就是上述的一切 0 与 1 之间的点，但端点不在内。在这个例子里，最小上界是点 1，它本身不在集内，但在它左边任意近的地方，却还有集里的点。我们给出下面定义，它显然包括上面 3 个例子的情况：

那个不一定属于点集的点 G，如果具有下面的性质，就称为点集

的最小上界：在 G 的右边，不再有点集的点，但是在 $G-\varepsilon$[①] 的右边，不管 ε 多小，总可以找到点集的至少一个点[②]。

以前说明了用无尽十进制数来表示实数的方法。在弄清楚关于最小上界概念之后，就可以利用该表示法来证明它的存在如下：

我们先看看哪些整数被点集的点越过，哪些不被越过。这是可能的，因为我们假定点集是完全确定了的。设 a 为被越过的最大整数，因而 $a+1$ 不被越过。这时就有两种可能性：或者 $a+1$ 被点集中的某个点所达到，或者不被任何点达到。在第一种情况下，$a+1$ 就是点集的最小上界。

在第二种情况下，我们把节 $a\cdots\overline{a+1}$ 平分为 10 段，得到以下各点

$$a=a.0, a+0.1=a.1, a+0.2=a.2, \cdots,$$
$$a+0.9=a.9, a+1=\overline{a+1.0}。$$

这里面十进制数的缩写是容易理解的。把这些数和点集里的点比较，就又从中找到一个被越过的最大数 $a.a_1$，$a.\overline{a_1+1}$($a_1=9$ 时，它就是 $a+1$)，它或者被达到，或者不被达到。

在第二种情况下，我们把节 $a.a_1\cdots a.\overline{a_1+1}$ 再平分为 10 段，得到

$$a.a_1 0, a.a_1 1, \cdots, a.a_1 9, a.\overline{a_1+1}0,$$

再把这些数和点集里的点相比较，就得到两个数 $a.a_1a_2$，$a.a_1\overline{a_2+1}$。我们看到，当这步骤继续下去时，或者有尽多步后得到一个属于点集的数，它就代表所求的最小上界；或者得到一个无尽序列的节，它们每一个含在前一个里面，而它们的长则无限制地趋于 0。

---

① 应明确规定 $\varepsilon>0$。——中译者

② 在 G 不属于点集的情况下，由于 ε 可以选得任意小，由上面定义可知，在节 $G-\varepsilon<x<G$ 里，有点集的无穷多个点，所以 G 是点集的聚点。反之，若 G 属于点集，则这个说法不一定正确，如上面第二例所示(孤立最小上界)。

根据魏尔斯特拉斯关于实数相等的定义，这些节的左端点和右端点确定同一个不尽十进制数①。这个数就代表问题中的最小上界。也许我可以把证明的思路简短地概括如下：

在证明最小上界的存在性时，是明确地把它用准确的十进制数代表的②。

由于第二个定理对函数论的重要意义，K·魏尔斯特拉斯在他的讲演里总是把它放在显著地位。因此，从第一定理的证明到第二定理，我们并没有中断讨论的一般性，我们只要指出，若把节的端点构成的序列 a 和 $a+1$ 看做无穷点集，则序列的极限点可以看做点集的聚点。用 x_0 表示聚点，则聚点的定义如下

若正数 ε 不论多小，$x_0-\varepsilon$ 和 $x_0+\varepsilon$ 之间总有点集的无穷多个点，则 x_0 称为点集的一个聚点。

证明第二定理时，我们采用的方法仍然是确实找到具有所需性质的 x_0 的十进制数表示。由此也可以看出，在每一个具体情况下，我们如何得到这样一个十进数来代表所求的实数。

首先仍把节 a 到 $a+1$③ 平分为 10 段

$$a.0,\ a.1,\ a.2,\ \cdots,\ a.9,\ \overline{a+1}.0,$$

在这些小段中，至少有一个含有点集的无穷多点。设这样一个节是 $a.a_1\cdots a.\overline{a_1+1}$。对于这个长度只有原节的 $\frac{1}{10}$ 的节，重复采用同样步骤，以得到 10 个长度相等的小节

$$a.a_1 0,\ a.a_1 1,\ \cdots,\ a.a_1 9,\ a.\overline{a_1+1}0,$$

这样得到的小节长度是最初的 $\frac{1}{100}$。它们中至少有一个节含有点集

① ［参看第一卷第 36—37 页（中译本第 26—27 页）］。

② ［与此类似，当然可以证明，有下界的点集必有最大下界。］

③ 在这里作者假设定理 2 中的 A 和 B 用实数 a 和 $a+1$ 代表，这只需适当地选取坐标原点、单位长和坐标轴正向即可。——中译者

的无穷多点。假定它是 $a.a_1a_2\cdots a.a_1\overline{a_2+1}$。你们看，我们又得到一个无尽序列的节，每一个含在前一个里，而它们的长趋于 0。这个节序列所决定的无尽十进制数 x_0 的确就代表所给点集的一个聚点，因为在序列的每一个节里，总有点集的无穷多点。所以，在这里又可以看出：我们证明聚点 x_0 的存在，是把 x_0 通过按一定规则进行的步骤，表示为十进制数，这样，我们就回到关于实数用十进制数代表的定义。

这里论述的关于点集的两个定理，向你们提供了例子，说明在单个自变数的讨论中，精确数学是怎样的一种事物，而且在证明中，我们如何运用了绝对准确性原理，而这个原理却是从直观引导出来的。

第二十三章

单变数 x 的函数 $y=f(x)$

§23.1 函数的抽象确定和经验确定(函数带概念)

到此为止,我们只谈论了自变数 x 的定义,现在我们要讨论函数 $y=f(x)$。我们要问,从精确数学的观点和从近似数学的观点,函数分别应当如何理解?

现代数学所采用的函数的最一般定义是从规定自变数 x 能有哪些值开始的。我们规定 x 在某个"点集"里变化,所以用的是几何的术语,但却是为康托-戴德金公理(第 9 页)所确定的,因而是在算术意义上来理解的。

若对应于点集里的每一个 x,有一个确定的 y,则 y 称为 x 的函数,用记号表示是 $y=f(x)$(这里 x 和 y 都是精确地确定的实数,即它们是十进制数,其中各位数都是确定的)。

通常 x 是在坐标轴的一段上连续地变动的,即它在两个固定点 m,n 之间的一切点所构成的点集上运动(用另一种说法,这样一个点集称为节 mn,而且当端点属于节时,称为闭节;当端点不属于节时,称为开节)。这样,我们就得到一个函数的老牌定义,例如 P·G·L·狄利克莱所用的定义:若对应于节里每一个数值 x,有一个完全确定的数值 y,y 就叫做 x 在那个节里的函数。

但这是精确数学里当时最广泛的函数概念。作为参考,我还要举出 A·普林斯海姆在《数学百科全书》里的文章:

Pringsheim, A.: *Allgemeine Funktionentheorie* (《一般函数

论》),*Enzyklop. d. math. Wiss.* 第 2 卷(1899 年结束)第 1—53 页。它以优美的方式说明函数概念是如何扩充的。

关于经验数学里的函数概念,我的想法已经在引言最后所举出的文章里说明了。下面接着所说的可以看做那里所介绍的概念的发挥。

M·帕什分两步论述了类似思想:他所讨论的范围虽窄,但很有趣味。第一步见 *Vorlesungen über neuere Geometrie*(《新几何讲义》,莱比锡,1882 年)[在那里,帕什从经验给定的点和直线开始,上升到抽象的表述①],第二步见 *Einleitung in die Differential und Integralrechnung*(《微积分引论》,莱比锡,1882 年)。我向你们特别推荐并请注意这本书;还可以参考帕什在 *Annalen* 第 30 卷,1887 年,第127—131 页②所作的与此平行的阐述。

最近丹麦数学家耶尔姆斯勒夫(J. Hjelmslev)在一系列著作中讨论了建立经验几何的问题,这种几何是按精确实验制约而与实践不矛盾的一门科学:

(1) 现实的几何。*Acta math.* 第 40 卷(1916 年),第 35—66 页。

(2) 自然几何。*Abhandl. aus dem math. Seminar der Hamburgischen Universität.* 第 2 卷(1923 年),第 1—36 页。

(3) 初等几何学(三卷),哥本哈根,1916—1921 年;这是丹麦一些学校采用的教材。

介绍了文献后,我们要问:

在经验几何里,我们应当怎样看待方程 $y = f(x)$?

我们先作一个直角坐标系,并提出这样的问题:当我们用钢笔

---

① [第二版有德恩所作的附录:Die Grundlegung der Geometrie in his-torischer Darstellung(《奠定几何基础的历史过程》,柏林,施普林格,1926 年)]

② P·杜·布瓦-雷蒙(du Bois-Reymond)在他的 *Allgemeinen Funktionen-theorie.*(《一般函数论》)中讨论了这个问题,而且是从"理想"和"经验"的区别中引出来的(他把空间直观作为经验),但他由此没有得到正面的数学好处,反而由于发挥了他的主观信念而迷失了方向。

或者用一种记录仪器(比方说,把气温作为时间的函数)画出一条连续曲线(图 23.1[①])的时候,作为 x 的函数,y 决定到什么程度。我们也可以不作出一条连续的线路而用一连串不连续的单个点来表示曲线(图 23.2)。从近处看,这一串的点的确是不连续的,但我们却感觉到那里有一条实在的曲线,而且从远处看,还会觉得看到一条连续曲线。

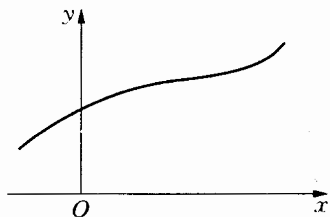


图 23.1

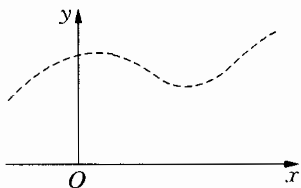


图 23.2

这里所举的简单例子,实际上是我们视觉的一个特性,这在别的场合也出现。例如当你看远处的树林时,你并没有感觉到它是由一棵棵的树构成的,你所得的印象是树林的轮廓;同样,当你看到拥挤的人群的头部时,也感到视觉的这个特征,即只是对人头的集体有一个概略印象。所以,我们要说,大量密切靠近的小事物,会给我们一个连续的印象。

上面提到用钢笔画出来的一条连续曲线。但是,如果用一个足够精密的显微镜去观察,就会发现,它是由许多小点点构成的。所以对上面所说的那句结论,还要补充如下:当我们认为面对着一个连续的事物的时候,如果仔细观察,就往往发现,它只是由紧密相接的小部分构成的。

上面只是把这个性质看成视觉对相互挨近的东西的印象。但是它也适用于在时间上紧密相接的视觉印象。在很短的时间间隔里一

① 为统一体例,书中的图按章编号,下同。——编辑

个挨一个的视察印象也会融合起来成为一个连续的东西。电影就是如此。

研究如何去解释这些现象，是生理学和心理学的事情。在这里，我只指出，我们的视网膜是由不连续的紧挨着排列的感锥构成的，很可能每单个的感锥只能留下某一个视觉印象。因此，对我们来说，观察到一个实际上连续的东西，或者一串充分接近的东西的集合，其间是没有区别的。当然，关于视觉的机制，关于视网膜的结构，关于光线对视网膜所起的化学变化，关于这些变化所经历的时间等等，生理学和心理学的研究已达到什么地步来揭示感觉的实质，我只好请专家们作出论断了。对这个问题，我们只能说明如上。

联系着上面的论述，我们现在问：一条经验曲线如何确定 y 作为 x 的函数？

我们回答：x 和 y 的确定都只能准确到一定的阀值。因此，经验曲线不是确定 y 作为 x 的精确的函数，对于每个 x，其对应的 y，只能确定到一定程度的准确度。

我们把这个事实写成

$$y = f(x) \pm \varepsilon,$$

其中不确定的 ε（在一定情况下，它也可以看成 x 的一个函数），满足一个不等式 $|\varepsilon| < \delta$，即这个小的数量 ε 是限在一个阀值 δ 之内的。

这样一个具有一定宽度的图形 $y = f(x) \pm \varepsilon$，在上述我在 1873 年的论文中叫做一个“函数带”。于是，我们就可以简短地说，由经验确定的曲线不是界定一个函数而是界定一个函数带。

§ 23.2 关于空间直观的引导作用

由于经验曲线只能确定函数到一定的准确度，现在我们把上面认为有争议的心理问题放在前沿地位。我们能否从空间直观的一条

曲线形象准确地感知一个函数?

这里所说的对曲线的“感知”,不能和下面的规律观点混淆:那就是,根据已假定为绝对准确的公理,精确地界定一条曲线,这里所涉及的毋宁说是关于一个具体图形的生动记录。

在前面已经说明如何在坐标轴上表示点集之后,我对上面提出的问题的答案就不会使人惊异了。在上面提到的 1873 年的论文中,是这样叙述的:即使一条曲线不是画出来的,也不是眼睛看到的,而只是“感知”的,它也只能有有限度的准确度,即它不属于精确数学的函数概念,而是联系着函数带概念。

下面我们将揭示这个观点的正确性:我将按照公理体系的要求,把多样化的曲线界说清楚,那样的曲线是根本不能感知的(通过这个方式,人们可以清楚地了解关于曲线的这些概念的产生规律)。

在数学文献中,已出现了与此相反的观点,特殊地。有 A·科普克(Köpke)在 *Math. Ann.* [①]第 29 卷(1887 年)第 123—140 页所提出的观点,我愿意在此讨论这个观点,因为虽然它不是准确地提出的,却传播得较广。

为了说明科普克的观点,我必须提到一些将来要详尽讨论的概念。

函数中有特殊的一类是连续函数,而连续函数中又有特殊的一类是可微函数。若函数可微任意多次,它就称为无尽次可微的。

科普克所说的是:我们所能准确感知的曲线 $y=f(x)$ 只限于 $f(x)$ 不但连续而且无尽次可微之款,至于那些曲线,其中的连续函数没有微商的,则根本不能感知。

你们看它和我的观点相去多远。我认为,从空间直观总是不准确的,而科普克则以为,空间直观足够确定一定类型的函数,但他又

① 查 *Math. Ann.* 里写的作者是 Alfred Köpcke in Ottensen。——中译者

排除了其他类的函数。

人们可能以为，生理学和心理学对这个问题已经研究过了。看来不是这样，在科普克和我之外，唯一详尽探讨这个问题的数学界人士是 M·帕什，他的观点见上面提到过的教材以及 *Math. Ann.* 第 30 卷(1887 年)。在不同科学接壤处，最重要问题的研究展开得太少了！与我们有关的可以指出这样一个被忽视的问题：现代的数学算术化(即把数学系统地建立在现代的实数概念基础上)，在数学界之外，了解的人越来越少，懂得它的重要性的人就更少了。

§23.3　自然规律的准确度(附关于物质构成的不同观点)

我愿意暂时离开主题而去论述自然哲学中与此联系着的下面问题：

当我们通过经验或者能从推理得到一个确切函数 $y=f(x)$ 时，它和自然规律的准确性有什么关系？它是精确的还是近似的？

让我们考虑物理或自然科学中所论及的任何一项规律，例如落体定律：

$$y=\frac{gt^2}{2}+ct+c'。$$

这个定律和观察吻合到什么程度呢？在这里，我们的答案仍然是：仅仅是近似的。同样，把一个物体投掷出去时，它的轨道也不是一条准确的抛物线，弹道曲线可以偏离抛物线很远。一个网球运行的曲线可能有向上的尖点。

那么如何看待一个作为绝对准确描述的规律和受到干扰所造成或观察中所产生的偏离呢？

这里有一条新的原则，即对自然说明的简约原则或思维经济

原则[①],它在目前的款可以叙述如下。

在陈述自然规律时,人们愿意(不是基于表述的考虑而是不自觉地)采取能足够准确地表述现象的最简洁的公式。

这里对落体的特款所说的话也适用于似乎是最确凿的自然定律,如质量与能量守恒定律。

在大多数(特别是化学方面)的研究中,质量基本上是作为不变的,用公式表示

$$M = 常数 + \varepsilon(t),$$

其中 $\varepsilon(t)$ 是时间的一个函数,但只有很小的值,此外,人们的经验对于它还没有任何确定的结论。

对于能量也是如此。在研究中,它基本上是作为不变的,但从绝对准确的观点来看,却绝非如此[②]。因此,我们必须说,我们最普遍而且对世界观最重要的自然规律也只是在有限的范围内和在有限度的准确度内为实验所验证。可以说:

用简洁公式作为自然规律是基于这样的愿望:把外界现象用可能最简单的工具去掌握。

从科学立场出发,对于每个自然规律——为了不把它看做教条,即不作为绝对正确而不必证明的定则——总要提出这样的问题:它在什么准确度的限度内为观察所验证?人们多数要发现,它的准确性比通常人们所设想的要小得多。

我想用一个例子来说明我所说的情况。这个例子采自天文学,一门准确性很高的经验科学。我利用了下面的著作:

Newcomb,S:*The elements of the four inner planets and the fundamental constants of Astronomy*(纽康:《4 个内行星的主要情

① 这是马赫主义者的用语。但在这里,作者强调了前提条件"能够足够准确地表述现象"。所以不能据此就认为作者是马赫主义者,请阅下文。——中译者

② 这个准确性问题至今考虑得极少,至少比质量守恒的研究要少很多。

况和天文学的基本常数》,华盛顿,1895 年)。

在这本著作里,纽康在他毕生研究大行星运动的基础上给出全部"天文常数"的最佳值,为当前天文界所采用;他还给出这些常数的确切的准确幅度①。

在第 118—120 页里,他讨论了牛顿关于两个质量之间的引力定律

$$f=\frac{kmm'}{r^2}\quad(k\text{ 为固定常量}),$$

在很大程度上为观测所证实。

他把观察区分为 3 类:

(1) 在地球上的观测,这时 r 所在的节是从若干厘米(实验室研究)到地球半径;

(2) 当 r 所在的节是从地球半径到日地距离时的观测(落体定律,月球和地球相对于太阳的运动);

(3) 当 r 大于日地距离 20 倍时的观测(远行星运动)。

关于牛顿引力定律的适用程度,或者说,这个定律为 3 类观测所完全证实的准确度,纽康指出,其不确定度为

(1) 在上述的第一个节里,等于引力 f 值的$\frac{1}{3}$;

(2) 对于地球半径到日地距离的节里,等于 f 值的 $\frac{1}{5\,000}$;

(3) 对于直到最远行星的节,不确定性极小。

当然,例如在第一类观测中,所说的准确度并不表明从定律算出的数值误差达到$\frac{1}{3}f$,而是说明现有的物质条件可能导致那样的误差。

---

① [晚近,在《数学百科全书》里有鲍申格尔(J. Bauschinger)的一篇综合性文章(Ⅵ$_2$, 17; 结束于 1919 年):"Bestimmung und Zusammenhang der astronomischen Konstanten, wo ebenfalls die Konstanten konsequent als Intervalle angegeben werden"("天文常数的测定及其间的关系,其中的常数也是一律用节的形式给出的。")]

我还要指出值得注意的一点：美国天文学家霍尔(Hall)为了解释水星运动的某些不规则性[1]，把牛顿定律用

$$f=\frac{kmm'}{r^{2+0.1574\cdot 10^{-7}}}$$

代替。

因此，无论如何，我们要说(这支持了我的论点)：

即使是经过实验证实为普遍适用的牛顿引力定律，其准确度也受很大限制。

与上面论述相联系，请让我谈谈有关的自然哲学问题。这是我在引言里所提到的，哥廷根哲学系关于本内克奖征文课题的专家报告里所讨论到的。

这里的问题也基本上是一些数量的近似估计和严格的实数定义的区别问题。特殊地，它涉及物质的内部结构问题(你们马上就要看到，这如何和我们所谈的较狭窄的内容相联系)。

关于物质的内部性质，有两种极端观点：

(1) 一种观点认为物质确实是连续地充满空间的(连续论)。

(2) 另一种观点认为空间是由分开的呈严格粒子状的质量积聚成的(极端原子理论)。

大多数科学家目前站在中间地位。他们看到物质是由分子构成的，但不是把分子看成点而是认为它本身就是一个复杂的世界。

第一个极端早些时特别是奥斯特瓦尔德(F. Ostwald)和他的学派(唯能论者(Energetiker))所主张的。他们排斥原子理论。第二种观点主要为法国经典学派的数学家所主张，特别是拉普拉斯(Laplace)，而且长期占统治地位；例如 1847 年亥姆霍兹

① [我们知道，爱因斯坦相对论迈进了一大步，并导致观点的改变。奥本海姆(S. Oppenheim)关于牛顿引力理论的详尽评论文章，见《百科全书》(Ⅵ$_2$，22，结束于 1920 年)。]

(Helmholtz)在他的最早著作里就把它作为指导观点。在那里，他假定互相分隔的质点通过中心力相互作用。这是从牛顿起在天文学中占统治地位观点的推广。

此外，这第二个极端有时候还在本质上更进了一步。人们提出这样的问题：时间是否也是不连续的？换句话说，会不会每一个运动都只是在不连续的时间间隔里跳跃地进行的？人们可以从电影引出这个想法，电影就是把不连续的过程表现为连续的。克利福德(W. K. Clifford)(*On theories of the physical forces*(《关于物理力的理论》), Proceedings of the Royal Institution 1870，重印在 *Lectures and Essays*(《讲演与论文》)第一卷，第 128—138 页，伦敦，1901 年)和玻尔兹曼(L. Boltzmann)实际上就是这样主张的①。

怎样能把如此不同的观点统一起来？这就是数学家所感兴趣的问题。我们必须说：

这些不同的假设可以归结为我们从感觉得到印象的不准确性；因为感觉的印象实际上不属于精确数学而属于近似数学。

我们将采取上述观点。但从根本上来说，这是严肃数学从事论证的前提。数学家既没有能力去判断"自然的本质"，也不以它为课题，他只能把它留给自然科学的专家们去做——如果真有可能得到确定结果的话。可以在适当的假设下，数学家能够而且应当判断，在多大程度下，不同的前提会导致大体相同的结果。对于近似数学，这的确是个大课题。

§ 23.4　经验曲线的属性：连通性、方向、曲率

在这些一般的论述以后，我回到一个确定的问题。

联系着经验曲线(例如用手画的曲线)每个人都要想到和它们有

① [人们知道上述关于不连续观点在现代量子论中获得了完全的发挥。]

关的某些性质。对于精确数学里的曲线 $y=f(x)$，能否把函数 $f(x)$ 加以限制，从而赋予它们类似的性质？

把我所想的性质列举如下：

(1) 经验曲线所具有的第一个性质是连续性，即曲线在它的最小局部是连通的。

这相当于 $f(x)$ 通过一切中间的值，或者更准确地：

若 $f(a)=A$，$f(b)=B$，则 $f(x)$ 在从 $x=a$ 到 $x=b$ 的节里，确实经过 A 与 B 之间的一切值。我们将要看到，所谓的连续函数有这个性质(图 23.3)。

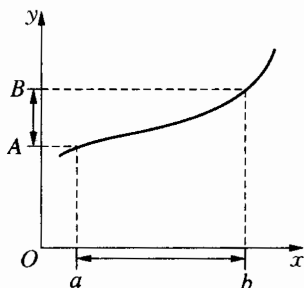


图 23.3

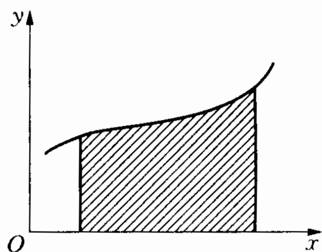


图 23.4

(2) 从经验曲线，我们看到第二个性质是，对应于它的连续性，它和 x 轴以及两条平行于 y 轴的直线包围着确定的面积(图 23.4)。在这里，我们假设曲线和每一条与 y 轴平行的直线交于一点。

面积存在这个要求并没有对连续函数 $f(x)$ 提出新的限制。我们将发现，对于在节 $a\leqslant x\leqslant b$ 连续的函数 $f(x)$，这个面积，即定积分 $\int_a^b f(x)\,\mathrm{d}x$ 是完全确定的。换句话说，每一个连续函数是可积的。

(3) 从经验曲线的一般路线来看，即从纵坐标的升降来看，我们发现，纵坐标在节内或节上有一个最大值和一个最小值，而且在

节内只有有尽多个极大值或极小值(图 23.5)。

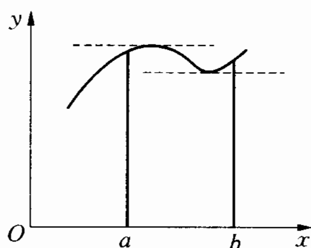


图 23.5

由算术方法确定的连续函数，在一个有穷节里，极大值和极小值可能有聚点。例如，若 $y=\sin\frac{1}{x}$(图23.6)，在从 $x=\frac{2}{\pi}$ 到 $x=-\frac{2}{\pi}$ 的节里，函数在 $y=+1$ 和 $y=-1$ 之间上下摆动无尽多次[在接近原点时，摆动次数无限制地增加]。或者，函数 $y=x\sin\frac{1}{x}$ 在一个含原点的有穷节里有无尽多个极大值和无尽多个极小值(图 23.7)。在第二例有一个优点，函数不但在原点的左右是连续的，它在原点本身也是连续的 ①。

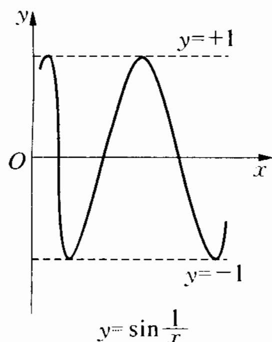


图 23.6

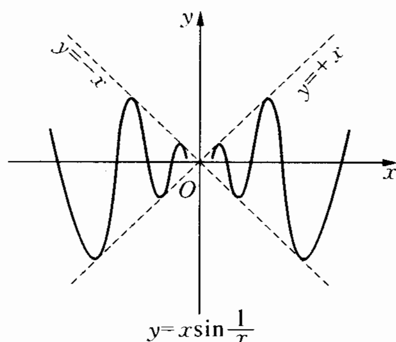


图 23.7

由此可见，由于不能排除一个连续函数 $y=f(x)$ 在一个有穷节里有无尽多个极大值和极小值，为了和经验曲线一致起来，我们还提出一个特殊的条件：

① 这里无形中规定了 $x=0$ 时，$\sin\frac{1}{x}=0$。这是可以接受的，因为虽然在 $x=0$ 处，$\sin\frac{1}{x}$ 没有确定值，但在任何别处，正弦的值是在 -1 和 $+1$ 之间。——中译者

函数 $y=f(x)$在所考虑的节里，可以分解成有尽多段，每段都是单调的。

(4) 对于经验曲线，我们赋予它一个方向。什么是方向？在这里我们给一个最自然的答案，它和我们从微积分学所知道的定义不能混同。

我们用一条有限宽度的小溪来表示所画的曲线。这样，所确定的纵坐标只准确到一个阈值 δ 的范围内（如果嫌小溪宽度太明显，下面我们就把它看成一条缩窄了的带）。为了确定方向，我们进行如下：

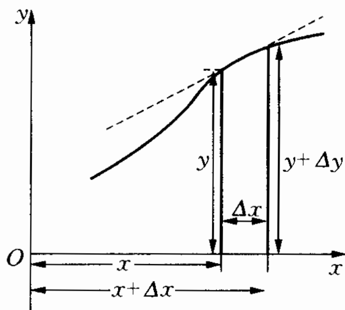


图 23.8

在坐标轴上取一段 Δx（图 23.8），相对于小溪（带）的宽，Δx 是大的，但相对于曲线的整体，它是小的。对于横标值 x 和 $x+\Delta x$，求出 y 和 $y+\Delta y$，并且把端点用一条直线连起来。这条线当然不能准确地决定，因为它所连的两点本身就不准确。但是，在实用上，它就代表那条带在那里的方向，它只有有限度的准确性。这个方向我们用差商$\dfrac{\Delta y}{\Delta x}$来度量①。这样，我们可以简单地说：

一条经验曲线在每处有一个方向，它用差商$\dfrac{\Delta y}{\Delta x}$表示，这个比值总是确定在一定的幅度里，这里的 $\Delta x\neq 0$ 是一个有一定大小的量。它相对于曲线整体是小的，相对于带的阈值是大的，经验曲线和从 x,y 到 $x+\Delta x,y+\Delta y$ 的连线实际上是近似地一致。

---

① 更一般地，可以令 $\dfrac{\Delta y}{\Delta x}=\dfrac{y_1-y_2}{x_1-x_2}$，其中 x_1 和 x_2 是在 x 的左和右一定距离处。

现在，另一方面，是否每一连续函数都有导数[①]?

导数的定义是

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}, \Delta x \neq 0,$$

并且写成$\frac{\mathrm{d}y}{\mathrm{d}x}$，其中 d 表示已经取了极限，但是在这里，我们必须考虑，从刚才所说的差商到导数要经过两个极限过程。这两个过程以一定次序实现：首先是带的宽，然后是坐标轴上线段 Δx 无限制地缩小。

关于极限过程，有一个普遍性的基本定理：先不管极限过程是怎样实现的，但当有不止一个极限过程的时候，它们的次序是必须明确规定的。

以前，数学家并不是根据这个基本定理行事的，他们相信每一个连续函数都有导数。但经过 B · 波尔查诺(B. Bolzano)、B · 黎曼和 K · 魏尔斯特拉斯等人的研究，人们知道连续函数可以根本没有确定的导数，即有时候$\frac{\Delta y}{\Delta x}$根本不存在像上面所界定的极限。我们以后将详尽地阐述这个问题。

相反的观点是怎样产生的呢？还不能肯定地说，毛病出在哪里。但是，看来那个相反的假定，认为一个连续函数必然有一个确定的导数，是由于人们设想经验曲线的宽度和增量 Δx 同时缩小，而且在缩小时，那条曲线和直线在节 Δx 里保持一致，而这正是它们形象的要点。一般地，对于严格的函数所代表的曲线的一小段同一条直线差不多。

如果这个解释是正确的，那么，错误结论之所以产生，就是由于我们同时而且同步地进行两个极限过程，可是这两个极限过程是应该一个接一个进行的。

---

① 亦称微商。——中译者

下面还要更准确地阐明连续函数可能没有导数。在这里，我们只需知道连续函数不一定有导数。因此，为了和经验曲线一致起来，既然经验曲线是有方向的，就必须对连续函数加一个条件，即它的导数存在。

(5) 同样没有问题的是，每条经验曲线有曲率。

一条曲线的曲率等于经过曲线上 3 个适当选取点的圆半径 ρ 的倒数$\frac{1}{\rho}$。若这 3 个点是

$$x,\ y;\ x+\Delta x,\ y+\Delta y;\ x+2\Delta x,\ y+2\Delta y+\Delta^2 y,$$

那么，曲率本质上要涉及二阶差分 $\Delta^2 y$，或者在除以$(\Delta x)^2$ 后，涉及二阶差商$\frac{\Delta^2 y}{(\Delta x)^2}$。它度量曲线和直线($\Delta^2 y=0$)间的偏离度，我们也可以把它写成

$$\frac{f(x+2\Delta x)-2f(x+\Delta x)+f(x)}{(\Delta x)^2}。$$

实际上我们如何选择这 3 点呢？显然和上面类似，它们的距离，相对于带的宽度是大的，相对于带的整体是小的。因为现在 3 个点都可以在带里移动，所以对于由经验给定的带的二阶差商(以及曲率)满足我们要求的精确度总是要小于一阶差商，以及曲线方向所具有的精确度。

当涉及三阶，四阶……差商的时候，这种不准确性还要增加。

为了保持和经验曲线类似，对于精确数学中的曲线，我们自然需要假定二阶或更高阶的导数存在。

把(4)，(5)两点合并，就有：

由经验给定的曲线有方向和曲率(一阶和高阶差商)这是人们感觉到的，为了使精确界定的函数所代表的曲线能有相应的性质，我们不但要假定该函数是连续的，在有穷节里有有尽多个极大值和极小值，还要明确假定它有一阶以及一系列更高阶导数(需用多少，就假

定多少)[①]。

§23.5 关于连续函数的柯西定义和经验曲线类似到什么程度?

上面最后的结论决定了我们下面的计划。为了对函数 $y=f(x)$ 所代表的曲线赋予从经验提炼出来的性质,按精确数学的观点,对函数 $f(x)$需有一系列严格的定义。

首先是连续性定义。

在这里,我按着历史顺序给出它以及有关的许多较初始的定义,这些都可以在下面两本基础性著作中找到:

Cauchy, A. L.: *Cours dánalyse*(《分析教程》第一卷,巴黎,1821年),以及 *Résumé des lecons données sur le calcul in finitésimal*(《微分学课程概要》),巴黎,1823 年。柯西(1789—1857 年)在综合技术学校讲授了这门课。

在这方面,另一位开拓者独立于柯西作出了努力,但他的名字只是在较晚才广泛地为人所知晓。他就是波尔查诺(1781—1848 年),他从 1817 年起就在布拉格工作。关于他对微积分的历史贡献,可参考斯托兹在 *Math. Ann.* 第 18 卷(1881 年)第 255—279 页的文章"Über Bolzanos Bedeutung für die Geschichte der Infinitesimalrechnung"("关于波尔查诺对微积分学的历史贡献")。在这里,和科学里常遇见的事是相同的:历史的发展总联系着一个人,但这种发展

---

① 为了一致性,我们当然可以作更多的规定。对于经验曲线,它在一点的方向和曲率等等也确定它在更大范围内的形状(这是显然的,因为我们是在更大范围内导出方向曲率等等的)。Ottensen 的 Otto 的科普克[②][我在第 20 页曾经提到他,在 *Hamburger Naturforscherversamm lung*(《汉堡科学研究文集》)对他有详尽的评论]对精确数学中的函数 $y=f(x)$,也要作那样的要求。这样,他得到的自然是比我们所讨论的更窄的一类函数。科普克要求这类函数比别的函数在直观上有更高的层次。我们的想法当然和他相反,我们认为,函数都不是直观的,只有函数带才是直观的。

② 应是 Alfred 的科普克,第 20 页译者注。——中译者

的可能性却是在许多地方同时出现的。

现在连续性的柯西定义是怎样的呢?

一个函数 $y=f(x)$在点 x_0 连续的条件有二：首先，它在这点唯一地确定；第二，对于无论多小的正数 η，总有一个具有如下性质的正数 ξ：在每一个含 x_0 在内的节里的两点 x, x'，只要 $|x-x'|<\xi$，就有 $|f(x)-f(x')|<\eta$。较直观地说：无论 η 如何选取，总有一个含 x_0 在内、长度为 $\xi>0$ 的这样一个节，在该节内部的一切点偶 x, x' 其函数值的差的绝对值小于 η。

若条件 $|f(x)-f(x')|<\eta$ 只对于以 x_0 为端点的节成立，则 $f(x)$称为在 x_0 半连续；而且按照长度为 ξ 的节是在 x_0 的右边或左边，$f(x)$依次称为右连续或左连续。现在，一个函数在一个节里连续意味着什么就容易界定了。这就是，它在节里每个点连续，但对于闭节，在它的左端点，只要求右连续；而在右端点，则只要求左连续。

在这些定义的基础上，我们现在证明：在闭节 $a \leqslant x \leqslant b$ 里，连续的函数经过 $f(a)$和 $f(b)$之间的一切值。

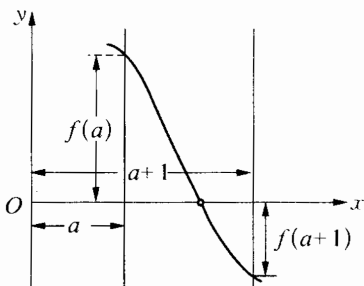


图 23.9

我们以一种特殊的形式证明这个定理，但这证明显示了一般原理。设 $f(x)$在整数 $x=a$ 和 $x=a+1$之间连续，并假定 $f(a)>0, f(a+1)<0$(图 23.9)。我们证明，$f(x)$在 a 和 $a+1$ 之间，至少有一次是零。

从经验来看，若一条连续曲线在 a 处位于横轴之上而在 $a+1$ 处位于横轴之下，则在 a 与 $a+1$ 之间显然要和横轴交于至少一点。可是理论上我们必须从所选取的连续性定义出发，在早已奠定的实数概念的基础上来证明这个定理。这里也是考验柯西关于连续性的定义：根据它可以证明由经验曲线的连续性(这是连续性的概念的

自然来源)立刻可以引出的定理。

重复一次所要证明的定理：按柯西定义在节 $a \leqslant x \leqslant a+1$ 连续的函数，若在 $x=a$ 处是正的，在 $x=a+1$ 处是负的，则在 a 和 $a+1$ 之间至少有一个零点。

证明的方式和关于点集的定理的证明一样，是“十进制证明”。

用十进制数

$$a_0.0,\ a_0.1,\ a_0.2,\ \cdots,\ a_0.9,\ \overline{a_0+1.0}$$

把 $a=a_0$ 到 a_0+1 那个节分为 10 等分。

若 $f(x)$ 在这些分点都不等于 0(若它在某分点是 0，就不再需要证明了)，那就有第一个节 $a_0.a_1 \cdots a_0.\overline{a_1+1}$，在它的左端，$f(x)$ 仍是正的，在它的右端，$f(x)$ 是负的。于是长为 $\frac{1}{10}$ 的这个节具有整个节所具有的情况。

把长度为 $\frac{1}{10}$ 的节 $a_0.a_1 \cdots a_0.\overline{a_1+1}$ 分为长度为 $\frac{1}{100}$ 的节

$$a_0.a_1,\ a_0.a_1 1,\ a_0.a_1 2,\ \cdots,\ a_0.a_1 9,\ a_0.\overline{a_1+1.0}。$$

在这些分点之一，$f(x)$ 可能是 0，我们假定不是这样。我们选取其中第一个这样的节，在它的左端，$f(x)$ 是正的，在它的右端 $f(x)$ 是负的。不断继续这个步骤，就得到一个节套，即一个无尽序列的节，每个都含在所有以前的节里，而它们的长度则无限制地向 0 减小。因此，这些节的左端点所对应的十进制数的序列和它的右端点所对应的序列有共同的十进制数 x_0 作为极限。我们要由柯西关于连续性的定义证明 $f(x_0)$ 等于 0。这个定理可以归结为一个一般性基本定理。

x_0 是一个递减的和一个递增的十进制数序列共同极限：对于前一序列，$f(x)>0$；对于后一序列，$f(x)<0$。这两个序列依次是

$$a_0,\ a_0.a_1,\ a_0.a_1 a_2,\cdots,$$

和

$$a_0+1,\ a_0.\overline{a_1+1},a_0.a_1\,\overline{a_2+1}。$$

于是对于连续函数，有以下定理：

若 x_0 是序列 $x_1,\ x_2,\cdots$ 的极限，则 $f(x_0)$ 也是函数值序列 $f(x_1),f(x_2),\cdots$ 的极限。

这定理从柯西关于连续性的定义以及极限值的定义可以简单地推得。

但据此，上述定理可以立即证明。因为 $f(x_0)$ 一方面是正数序列的极限；另一方面又是负数序列的极限，则只有当 $f(x_0)=0$ 时，这两方面才不出现矛盾。

在证明中，我们最后用了一个比所证定理几乎更为重要的定理。用记号表示，这就是：若 $f(x)$ 连续，则

$$f(\lim x)=\lim f(x)。$$

由这个定理可以推证另一个以后有用的定理：

设 x 为横轴上有理点 $x=\frac{m}{n}$ 的集合。从它们可以确定一切无理数，即把无理数作为无尽有理数序列的极限。因此，对于无理数 x_0

$$f(x_0)=f(\lim \text{有理数}\ x)=\lim f(\text{有理数}\ x)。$$

用文字叙述，即：若一个连续函数在一个节里的一切有理点的值已经确定，则它在该节里一切点的值也就确定。若回忆已经论述过的关于在一个节 $(a,\ b)$ 里处处稠密的点集的概念，则上述结论即可推广。处处稠密点集的特点是：在 $(a,\ b)$ 节的每一个无论多小的子节里，总有一个该点集的点，我们可以看出，$(a,\ b)$ 节里的点都是聚点，即对于节中每点 x_0，即使不属于该点集，也有属于点集的点序列，以 x_0 为极限值。于是得到定理：若连续函数的值在一个节里的一个处处稠密的点集的点已知，则它在该节里的一切

点都已确定。

作为处处稠密的点集，可以选取一切有尽十进制数（因为它们在横轴上已构成一个处处稠密集），这样就有系理：

若一个连续函数在一切用十进制数代表的点的值已知，则函数完全确定。

§ 23.6　连续函数的可积性

现在到了第二点，即关于曲线下面的面积问题。

我们感到每一条经验曲线段和 x 轴以及经过它的两个端点的纵线包围一定的面积。为了近似地处理这个问题，可以把纸上图形剪下来称一下，然后和一个已知面积，例如 100 平方米的纸的重量比较，这样往往能得到可用的结果。

另一种实用方法是机械求积法或数值求积法，用的是梯形公式和辛普森（Simpson）定则。现只谈梯形公式。把所给节分成充分小的段 Δx（可以让它们长度相等），使各个 Δx 上的曲线段可以近似地看做直线段，因而这部分的面积可以作为梯形面积计算（图 23.10）。把一切梯形面积相加，就得到曲线下的面积

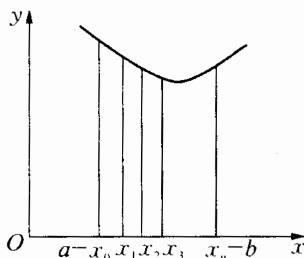


图 23.10

$$\Delta x\,\frac{y_0+y_1}{2}+\Delta x\,\frac{y_1+y_2}{2}+\cdots+\Delta x\,\frac{y_{n-1}+y_n}{2}$$

$$=\frac{\Delta x}{2}(y_0+2y_1+\cdots+2y_{n-1}+y_n)。$$

第三种实用方法是用平面求积仪①。

① ［参考第二卷第 11—16 页（中译本第 12—17 页）。］

在运用实际方法确定面积时，主要的误差来源，一般地不在于把那些矩形放在一起来代替那块面积，也不在于 Δx 选择的任意性；最重要的还是作为那块面积边界所画出，或者用其他经验方法所给出的曲线的有限宽度，例如一条 10 厘米长，$\frac{1}{3}$毫米宽的带，由于宽度所产生的误差是$\frac{1}{3}$平方厘米。

那么我们如何在理论上把握这个事物呢？

理论上的论述从总和 $S=\sum_{1}^{n}\Delta x_{\mu}\cdot f_{\mu}(x)$ 开始，其中 n 是把 $a\leqslant x\leqslant b$节任意细分中的子节数目，而 $f_{\mu}(x)$表示子节 Δx_{n} 中的任意一个纵坐标。我们考察，当子节的最大长度 Δx_{μ}①趋于 0，因而 n 自然无限制地增长时，这个总和的变化。若对于满足条件 $\Delta x_{\mu}\to 0$ 的每一个细分的序列和每一种对于 $f_{\mu}(x)$的选择，这个总和都有同一个极限，这极限就用 $\int_{a}^{b}f(x)\mathrm{d}x$ 表示，并且称为 $f(x)$在下限 a 与上限 b 之间的定积分。下面要判断的问题是：

这样一个极限什么时候存在？即函数 $f(x)$什么时候可积？

第一个回答这个问题并且以一般方式加以处理的是黎曼，在他的就职论文(1854 年)里。其论文是："Über die Darstellbarkeit einer willkürlichen Funktion durch eine trigonometrische Reihe"②("关于用三角级数代表任意函数的可能性问题")。

我们从考察在 $a\leqslant x\leqslant b$ 节有界的函数开始，在这里，绝对值不超过一个固定数 K 的函数称为有界。这样一个函数不一定连续。我们在下面将要证明，它在整个节内以及在每个子节内有一个最小上界和一个最大下界。G_{μ} 是函数在子节 Δx_{μ} 上的最小上界，g_{μ} 是最大下界，则 $f_{\mu}(x)$ 就介乎 g_{μ} 和G_{μ} 之间(图 23.11)。于是我们得：我们

① 原为 Δ_{n}，似为 Δx_{μ}，下同。——编辑

② *Mathematische Werke*(《数学文集》)，第 213—251 页。

原来的总和 $\sum \Delta x_\mu f_\mu(x)$ 的值介于上界 $\sum \Delta x_\mu G_\mu$ 和下界 $\sum \Delta x_\mu g_\mu$ 之间，而这两和之差是 $\sum \Delta x_\mu D_\mu$，其中 $D_\mu = G_\mu - g_\mu$，并且称为 $f(x)$ 在第 μ 个小节上的变差。

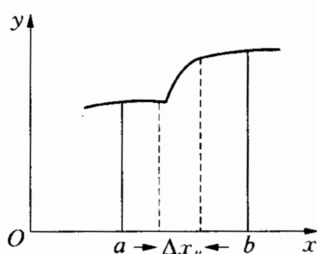


图 23.11

可以证明(但我们不加详述)，对于有界函数，当 $\Delta x_\mu \to 0$ 时，$\sum \Delta x_\mu G_\mu$ 和 $\sum \Delta x_\mu g_\mu$ 趋于两个极限 A 和 B，而 $\sum f(x_\mu)\Delta x_\mu$ 也总在它们之间，若极限 A, B 都等于 C，则 $\sum f(x_\mu)\Delta x_\mu$ 的极限必然存在，并且也等于 C，可是，如果在子节的最大长度 Δx_μ 充分小时，$\sum \Delta x_\mu D_\mu$ 可以变得任意小，则 A, B 肯定相等，因而有界函数 $f(x)$ 可积。

特殊地，我们要证明，对于每一个在闭节 $a \leqslant x \leqslant b$ 之间连续的函数——这样一个函数必然有界——可以使和 $\sum \Delta x_\mu D_\mu$ 任意小。

在我进行证明之前，我还必须指出关于连续性的一个细节，但我不准备详细讨论。

上面已经指出：若对于每一个 ξ，使得对于一切满足 $|x-x_0|<\xi$ 的 x，就有 $|f(x)-f(x_0)|<\eta$，则函数 $f(x)$ 称为在 x_0 处连续。

现在，当 x_0 在 $a \cdots b$ 节移动时，连续性的变化如何？即当 x_0 在节上运动时，连续性的程度[①]会有什么变化？

当一条曲线这里平些，在那里陡些时，就出现这个问题。显然，对于同一个 η，在较平的地方 x_{01}，ξ 可以大些；在较陡的地方 x_{02}，ξ 就要小些。所以，已经给定 η，在不同的 x_0［按照曲线在那里较平或较

① 连续性条件中 ξ 的值体现连续性程度。——中译者

陡(图 23.12)],ξ 的值可能区别很大。

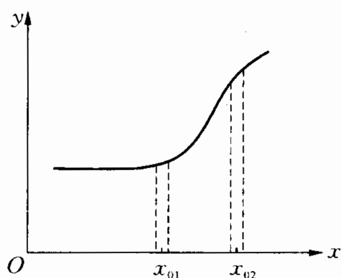


图 23.12

可以利用一条辅助曲线来清楚地表现连续程度。

选择 x_0 为横坐标,$10^4\xi_0$ 为纵坐标。所得的曲线(对于已给的 η,它表现 ξ 依赖于 x_0 的情况)在整个节内,其纵坐标都是正的。但是还不能排除辅助曲线会有图 23.13 所示的情况。因为纵标在一个或更多地方会任意接近 0。而由于 $f(x)$ 的连续性,曲线的纵坐标在那样一个地方的旁边,必须突然跳到高处,这样才能满足 ξ 到处有正值的要求。于是辅助曲线将在那些地方有不连续点。若发生这种情况,原来的函数在节里就称为"非一致连续",我们说:

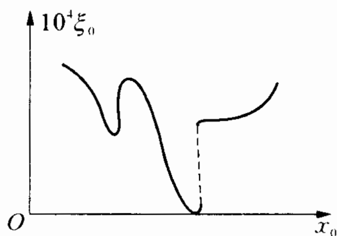


图 23.13

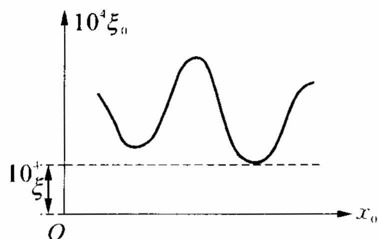


图 23.14

假定了函数在节的每个单独的点连续并不能排除在整个节上,它的连续性是非一致的,即不能排除,有这么一些地方,在那里,ξ 的值比任何已给的值都小。

与此相反,若曲线虽然上下摆动,但不出现上述那种例外的情况(图 23.14),函数在节里就是一致连续的,即对于节里一切点偶 x,x_0,只要 $|x-x_0|<\xi$,就同时有 $|f(x)-f(x_0)|<\eta$。

现在遇到这样一个事实:为简明起见,我们略去它的证明。可

以证明，对于在闭节里连续的函数，上述设想的非一致连续性的不良情况根本不可能出现，因为闭节是自紧的[①]，即它的每一个点序列都有一个属于节的聚点。

于是有定理：在闭节里连续的函数，也在该节里一致连续。

详情可参考 A·普林斯海姆：*Allgemeine Funktionentheorie*（《一般函数论》）。Enzyklopàdie der math. Wissenschaften，第 2 卷第 18 页[②]。

在这里，我如此详细地提请注意非一致连续性的可能性，是因为在后面讨论级数收敛问题还要遇到类似的可能性。

现在我们回到连续函数的可积性问题。

因为知道 $f(x)$ 在闭节 $a \leqslant x \leqslant b$ 一致连续，我们可以一劳永逸地找到一个 ξ，使得在整个节里，只要 $|x-x_0|<\xi$，就有 $|f(x)-f(x_0)|<\eta$。因此，若选择长度 $\Delta x_\mu<\xi$，则在节 Δx_μ 里，变差 $G_\mu-g_\mu<\eta$，即 $D_\mu<\eta$。对于 Δx_μ 和变差的积之和，现在有

$$\sum \Delta x_\mu D_\mu<\eta \sum \Delta x_\mu=\eta(b-a),$$

即：对于在闭节连续的函数 $f(x)$，我们可以使 $\sum \Delta x_\mu D_\mu$ 小于 $\eta(b-a)$，因而小于任意已给的正数。可见这个函数总是可积的。

§ 23.7 关于最大值和最小值的存在定理

我们现在考虑第三点：关于连续函数的最大值和最小值。

一条经验曲线的纵坐标在一个节里总有一个最小的和一个最大的值[在某些情况下，它们会出现在节的端点（图 23.15）]。问题是，在

① in sich kompact. ——中译者

② [还可以参考 A·普林斯海姆的 *vorlesungen über Zahlen-und Funktionenlehre*（《实数与函数理论讲义》），II. 1，莱比锡，1925 年，第 54—56 页和 F·豪斯多夫（Hausdorff）的 *Grundzüge der Mengenlehre*（《集合论基础》），第二版，柏林，1927 年，第 8 章，第 197 页。]

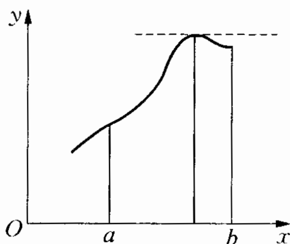


图 23.15

精确数学的意义下一条理想曲线，或说，一个函数 $y=f(x)$（它在闭节里自然是有界的），在$a\cdots b$里是否有个最大值？

在这里，我们必须区分最小上界和最大值。对于最小上界，首先有定理：若 $f(x)$ 在 $a\leqslant x\leqslant b$ 节里有界，节里的 y 值就必有一个最小上界（最大下界），这是在第 36 页里指出过的。

定理的证明可以很容易从我们关于点集的第一个定理（第 11，12 页）推出。实际上，对于节里每一个 x，总有一个确定的 y，所以在 y 轴上，对应于 x 轴的节，有一个完全确定的有上界的点集。但是，对于这样一个点集，我们已经证明最小上界存在。

但是还不知道的是，该点集的最小上界是否属于该点集，即它是否是对应于一定的 x 的纵坐标，或者说，最小上界可否达到。我要先举一个例子来说明，情况可能是否定的。设用一个半圆来界定 $f(x)$在$+1$ 和-1 之间的一切点，半圆圆心在原点，半径等于 1，只是对于$x=0$，设 $f(x)=0$，于是

$x\neq 0$ 时，$y=|\sqrt{1-x^2}|$；

$x=0$ 时，$y=0$。

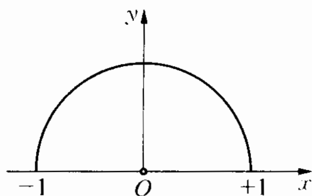


图 23.16

如图 23.16 所示，这个函数是完全界定的，却不是连续的；因为它唯一能达到最小上界的地方是 $x=0$，但根据定义，$f(0)=0$。

与此相反，在节 $a\leqslant x\leqslant b$ 里连续的函数必达到它的最小上界，即在节里连续函数有一个最大值。定理成立的关键在于节的端点属于节。K•魏尔斯特拉斯在他的讲义里总是把这个定理放在突出的地位。

首先，在闭节里连续的函数是有上界的，因而有最小上界 G。现

在利用十进位数的步骤,可以得到结论:存在一个 x_0,具有所需要的以下性质:在 x_0 的无论多小的邻域里,G 是 $f(x)$的上界。我们把这个节分成 10 等份,选择第一个以 G 为最小上界的子节,并对于这个子节,我们把它像原先的节那样处理。把这个步骤无限制地继续下去,就得到一个节套,它确定所需要的 x_0。虽然对于每个有上界的函数 f 得到了上述结论,但 $f(x_0)=G$ 的结论,却不是对每个有界函数都正确的,可是它对于在闭节里连续的每个函数,则是正确的。我们将证明后一个结论。

我们用反证法。假定 $f(x_0)=G'$, G'小于G。由于函数 $f(x)$的连续性,有一个含 x_0 的节,使得

$$|f(x)-f(x_0)|\leqslant\eta,$$

其中 η 选得小于 $G-G'$。这样,在含 x_0 的那个节里,函数值就在 $G'-\eta$ 和 $G'+\eta$ 之间。但由于 $G'+\eta$ 总是小于 G,对于那个节,G 不是$f(x)$的最小上界;可是根据 x_0 的定义,在那个节里,$f(x)$又必须以 G 为最小上界。这个矛盾可以叙述如下:若 $f(x_0)=G'<G$,则由于$f(x)$在 x_0。邻近的连续性,总有一个含 x_0 在内的节,在那里一切函数值比 G 至少小于一个固定的正值 δ,这和假设在每一个含 x_0 的节里 G 是最小上界是不相容的。

归纳一下,我们看到,经验曲线的连续性重现于精确数学中,在闭节里连续的函数,它的曲线也具有以下性质:包围一个面积,而且最大值和最小值存在。

如已指出的,为了和经验函数一致,我们还要明确地假定函数 $f(x)$在每一个有限节里只有有尽多个极大值和极小值,这也就是说,在每一个足够小的邻域里,有一个最大值和一个最小值。

§ 23.8　4 个广义导数

$f(x)$的导数存在性的问题是怎样的呢?

我们首先说，对于经验曲线，我们看到，方向是用差商$\frac{\Delta y}{\Delta x}$($\Delta x$在一定的数量级内)界定的。对于精确数学中的函数，方向则用导数

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}。$$

界定。

问题是：这样一个极限是否存在？也就是：对于一切趋于零的序列Δx_ν，极限是否一样？

为了把这个问题完全弄清楚，我们从一个例子开始。

取上面已经提到过的函数(图 23.7)

$$y=x\sin\frac{1}{x}。$$

它在原点左右都在直线$y=+x$和$y=-x$之间摆动，而且越接近原点，摆动得越密；在原点本身，它却是连续的(参考柯西的定义)。对坐标原点和任意另一个点(x, y)作差商，显然

$$\frac{\Delta y}{\Delta x}=\frac{y}{x}=\sin\frac{1}{x}。$$

现在对于导数，我们能说些什么呢？

函数$\sin\frac{1}{x}$对于无限制变小的x，在$+1$与-1之间无限频繁地上下摆动，因此，当$\Delta x\to 0$时，差商$\frac{\Delta y}{\Delta x}$在$+1$与$-1$之间摆动，但不趋于一个固定的极限。无论把$\Delta x$取得多么小，差商的值所构成的集合的最小上界和最大下界是$+1$和-1。我们看到，序列$\frac{\Delta y_\nu}{\Delta x_\nu}$的极限的存在及其大小依赖于序列$\Delta x_\nu$的选择。若选序列$\Delta x_\nu$，使得极限存在，则极限值可以是从$-1$到$+1$间的任意值。对于正值$\delta$，用$G(0, \delta)$表示在零点旁的节$0<x<\delta$里一切$\frac{\Delta y}{\Delta x}$的最小上界，$g(0, \delta)$表示最

大下界，再用 D^+ 和 D_+ 表示；$\lim\limits_{\delta\to 0}G(0,\ \delta)$，$\lim\limits_{\delta\to 0}g(0,\ \delta)$，然后对于负值 δ，用 D^- 与 D_- 表示相应的极限值，则在我们的例子里，有

$$D^+=+1,\ D_+=-1,\ D^-=+1,\ D_-=-1。$$

这样界定的 4 个数称为函数 $y=x\sin\dfrac{1}{x}$ 在 0 处的广义导数。特殊地，D^+，D_+ 称为右方上导数和右方下导数，D^- 和 D_- 称为左方上导数和左方下导数。从下面的论述可以看到，对于每一个连续函数，都可以有这样 4 个导数。

为了便于直观，我们把上面的例子略微改动，使函数在 y 轴的左右不再对称。令：

在原点右方，$y=ax\sin\dfrac{1}{x}$，

在原点左方，$y=bx\sin\dfrac{1}{x}$，

其中 $a+b\neq 0$①。于是，它的几何图形是一条曲线，右方在 $+ax$ 和 $-ax$ 之间，左方在 $+bx$ 和 $-bx$ 之间摆动，但在原点是连续的（图 23.17）。

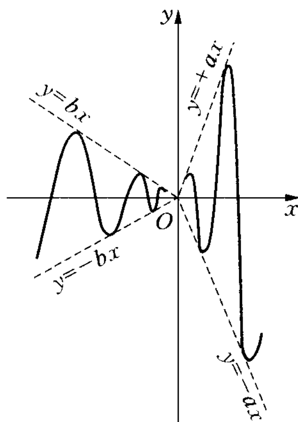


图 23.17

这样，对于无限制变小的 Δx，4 个差商的极限就是

$D^+=a,\ D_+=-a,$

$D^-=b,\ D_-=-b,$

就是说：在这个例子里，4 个导数彼此都不相同。

我们可以向普遍性再进一步。上述的例子还有一点特殊性。当 Δx 趋于零时，差商 $\dfrac{\Delta y}{\Delta x}$ 的值本身就无尽多次达到其最小上界和最大下界。但我们容易举出与此不同的例子。

① 参照图 23.17，应假设 $a>0$，$b>0$，$a\neq b$。——中译者

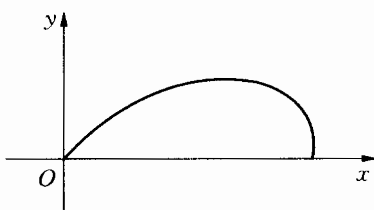


图 23.18

设 $y=y_1\sin\frac{1}{x}$，其中 y_1 是一条曲线的纵坐标，这条曲线就像图 23.18 那样，经过原点时和 x 轴成 45°角。

这时，$\frac{\Delta y}{\Delta x}$ 达不到它的最小上界+1 和最大下界−1，但能任意接近它们。

为了有显式表示，例如选取 y_1，使上述曲线就是双纽线，它的方程是

$$(x^2+y_1^2)^2=x^2-y_1^2。$$

于是在第一象限里的一段就是

$$y_1=\left|\sqrt{\frac{|\sqrt{8x^2+1}|-(2x^2+1)}{2}}\right|。$$

这样，我们的函数就是

$$y=\left|\sqrt{\frac{|\sqrt{8x^2+1}|-(2x^2+1)}{2}}\right|\sin\frac{1}{x}。$$

当 $x>0$ 时，对于 0,0 和 x，y，差商

$$\frac{\Delta y}{\Delta x}=\frac{1}{x}\left|\sqrt{\frac{|\sqrt{8x^2+1}|-(2x^2+1)}{2}}\right|\sin\frac{1}{x}$$

$$=\left|\sqrt{\frac{2(1-x^2)}{|\sqrt{8x^2+1}|+(2x^2+1)}}\right|\sin\frac{1}{x}。$$

当 $x\to 0$ 时，第一个因子无限制地接近 1，而第二个因子则在+1 与−1之间上下摆动无尽多次。这样，当 Δx 趋于零时，$\frac{\Delta y}{\Delta x}$ 在 $D^+=+1$ 和 $D^-=-1$ 之间摆动，但总达不到这两个值。

通过这些例子，下面所说就清楚了：设已给任意连续函数 $y=f(x)$。取点(x, y)，并向右方取 Δx 及其对应的Δy，则当 Δx 无限制地减小时，$\frac{\Delta y}{\Delta x}$ 就有两个极限值：最小上界 D^{+} 和最大下界 D_{+}。同样，当 Δx 从左方趋于零时，$\frac{\Delta y}{\Delta x}$ 有两个确定的极限值D^{-} 和 D_{-}。在这里，$+\infty$和$-\infty$也可能作为极限值出现。因此，不管连续函数 $f(x)$的具体情况如何，当 Δx 趋于零时，$\frac{\Delta y}{\Delta x}$ 总有 4 个极限值，从这边是 D^{+} 和 D_{+}，从另一边是 D^{-} 和 D_{-}①。

当 $D^{+}=D_{+}$ 时，我们说，函数有"右导数"；当 $D^{-}=D_{-}$ 时，我们说，它有左导数。

当右方上、下导数和左方上、下导数都相等时，而且只在此时，我们说，函数在那里有唯一确定的导数。

当我们按照这种方式弄清楚函数在每一点有导数需满足几个条件时，我们必须问，这样的函数根本存在否。这个问题是极端对立的两个问题之一，另一个极端来自不假思索的习惯观点，即每一个连续函数都该有导数②。

---

① ［关于 4 个导数更详尽的探讨，见 C · 卡拉西奥多里(Carathéodory)：*Vorlesungen über reelle Funktionen*(《实变函数讲义》，第二版，莱比锡，1927 年)。］

② 在佩林(Jean Perrin)的著名著作 *Les Atomes*(《原子》)［由洛特莫泽(A. Lottermoser)译成德文(1914 年)］里，有与此相关的提示。他在该书的前言第Ⅸ页和第Ⅹ页中说："利用显微镜观察布朗运动，看到浮在溶液里的每个小颗粒来回运动时，我们被实验的真实完全吸引住了。为了对它的轨迹作切线，需要作连接轨迹上非常接近的两点(即该小颗粒在两个邻近时刻的位置)的直线，并找出它的极限位置。可是，在我们的研究过程中，两次观察的时刻间隔无论多小，这直线的方向总是持续地变化着。由这项研究，一个不带成见的观察者只能得到函数没有导数的感觉，而丝毫得不到有切线的曲线的感觉。"

还可以参阅 E · 波雷尔(Borel)在他 1912 年的讲演"Les théories moléculaires et les mathématiques"("分子理论及其数学")中的有趣评述。该讲演在他的 *Introduction géométrique à quelques théories physiques*(《一些物理理论的几何导引》，巴黎，1914 年)中作为注记Ⅶ重印了。

§23.9　魏尔斯特拉斯不可微函数；它的形象概述

在上述一般论述之后，我们转而详细地讨论这样一个例子：一个连续函数在任何处没有符合我们定义的导数。

我们选取魏尔斯特拉斯的著名函数。他大约在 1861 年找到这函数，可是直到 1874 年在 P·杜·布瓦-雷蒙的文章里才附带地公开发表(*J. f. Math.* 第 79 卷，1875 年)①。

魏尔斯特拉斯函数是用具有形式

$$y=\sum_{0}^{\infty} b^{\nu}\cos a^{\nu}\pi x$$

的不尽三角级数确定的，其中 $b>0$，但是要使级数收敛，它必须<1，而且 a 和乘积 ab 还需满足一些需要进一步给定的条件。

我们首先通过一个数值的例子来说明函数的构成。设 $b=\frac{1}{2}$，$a=5$，则

$$y=\cos\pi x+\frac{1}{2}\cos 5\pi x+\frac{1}{4}\cos 25\pi x+\frac{1}{8}\cos 125\pi x+\cdots$$

---

① [魏尔斯特拉斯本人对该函数论述见他的文集第 2 卷第 71—74 页。此外，可以阅读魏尔斯特拉斯致 P·杜·布瓦-雷蒙和柯尼希斯伯格(L. Koenigsberger)的非常有趣的信。这两封信在 *Acta Math.* 第 39 卷(1932 年)第 199—239 页公开发表。在魏尔斯特拉斯之前 30 年，波尔查诺已经构造了无处可微连续函数的例，但直到几年前才被发现。对此，可以参考 G. Kowalewski："Bolzanos Verfahren zur Herstellung einer nirgends differenzierbaren stetigen Funktion"(柯瓦莱夫斯基："波尔查诺建立无处可微连续函数的方法")，Leipziger Ber. (*Math. Phys.*)第 74 卷(1923 年)，第 315—319 页。关于这里所谈的问题的其他文献还可以指出 K·克诺普："Ein einfaches Verfahren zur Bildung stetiger nirgends differenzierbarer Funktionen"("构造无处可微连续函数的一个简单方法")，*Math. Zechr.* 第 2 卷(1918 年)，第 1—26 页和 A. Rósenthal："Neuere Untersuchungen über Funktionen reeller Veränderlichen"(罗森塔尔："关于实变函数的新研究")，*Enzyklopädie der Math. W.* 第 2 卷，第 1091—1096 页。]

我们先考察部分曲线

$$y_0=\cos\pi x,$$

$$y_1=\frac{1}{2}\cos 5\pi x,$$

$$y_2=\frac{1}{4}\cos 25\pi x,$$

$$\cdots\cdots$$

把它们摞起来，就得到曲线 y。

曲线 $y_0=\cos\pi x$ 是一条通常的余弦曲线，半波长是 1，第一个正零点位于 $x=\frac{1}{2}$。

函数 $y_1=\frac{1}{2}\cos 5\pi x$ 确定一条波状曲线，它在 $+\frac{1}{2}$ 和 $-\frac{1}{2}$ 之间摆动，第一个正零点位于 $x=\frac{1}{10}$，它的半波长是 $\frac{1}{5}$。所以波状曲线 y_1 比 y_0 陡。这里，部分曲线的陡度用它在一个零点处的斜率绝对值来衡量。这样，对于 y_0，陡度 $\left|\frac{\mathrm{d}y_0}{\mathrm{d}x}\right|_{x=\frac{1}{2}}=\pi$，对于 y_1，它增加到 $\left|\frac{\mathrm{d}y_1}{\mathrm{d}x}\right|_{x=\frac{1}{10}}=\frac{5\pi}{2}$。

函数 $y_2=\frac{1}{4}\cos 25\pi x$ 确定一个波状曲线，它比前一个更陡。第一个正零点位于 $x=\frac{1}{50}$，半波长是 $\frac{1}{25}$，纵坐标在 $+\frac{1}{4}$ 到 $-\frac{1}{4}$ 之间摆动，陡度是 $\frac{25\pi}{4}$。

在那以下的部分曲线，一个比一个陡，高度一个比一个小，它的幅度按着比值为 $\frac{1}{2}$ 的几何序列减小，但波长比幅度减小得快得多，即按着比值为 $\frac{1}{5}$ 的几何序列减小，而陡度则增加得非常迅猛，按照比值

为$\frac{5}{2}$的几何序列增加。由这些观察，我们指出对下文具有重要意义的一个关键结论：用以摞成最后曲线的单个部分曲线是波状曲线，它们的幅度逐渐减小，但陡度非常迅猛地增加。

现在我们离开数值的例，回到原先的函数$y=\sum_{0}^{\infty} b^{\nu}\cos a^{\nu}\pi x$，对于它，我们区别

(1) 部分曲线　$y_{\nu}=b^{\nu}\cos a^{\nu}\pi x$;

(2) 近似曲线　$Y_{m}=\sum_{0}^{m} b^{\nu}\cos a^{\nu}\pi x$

(Y_m 是前 $m+1$ 条部分曲线之和)。

当$a>1, b<1$时，对于部分曲线总有定理：当ν增加时，部分曲线的幅度和波长无限制地减小。可是部分曲线是否像我们上面所举的具体的数值的例子那样越来越陡呢？

微商$\frac{\mathrm{d}y_{\nu}}{\mathrm{d}x}$等于$-a^{\nu}b^{\nu}\pi\sin a^{\nu}\pi x$；因此，在$y_{\nu}$的一个零点

$$\left|\frac{\mathrm{d}y_{\nu}}{\mathrm{d}x}\right|=(ab)^{\nu}\pi。$$

所以，只要$ab>1$，当ν无限制地增大时，尽管部分曲线的高度无限制地减小，它的陡度却增加。

我们现在略谈近似曲线。只有当我们弄清楚这样一条近似曲线如何从前一条得到后，我们才能够了解，当这个步骤无限制地继续下去时，所得到的最后曲线的某些形象。

首先有

$$Y_{1}=\cos\pi x+b\cos a\pi x,$$

因而我们可以想象，把一条较细密的波状线加到普通的余弦曲线上，就可以得到近似曲线Y_{1}。

其次

$$Y_2 = \cos \pi x + b\cos a\pi x + b^2 \cos a^2 \pi x,$$

因而在这里，对于前一条曲线加上了一条更细密的波状线。

如果说，我们在这里给出近似曲线的有机构成的话①，那么，当我们想要获得所求极限曲线的整个轮廓时，我们的直观能力就迅速变弱了，我们宁可满足于迅速地转入下面那种逻辑分析：

图 23.19 所示是魏尔斯特拉斯曲线 $\sum\limits_{0}^{\infty} b^{\nu}\cos a^{\nu}\pi x\left(b = \frac{1}{2}, a = 5\right)$ 的近似曲线。

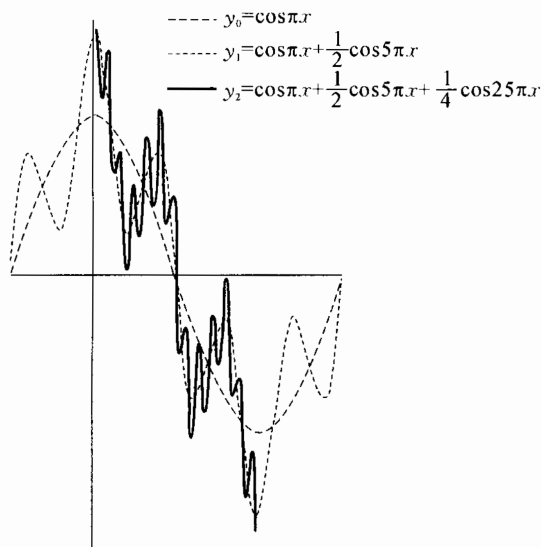


图 23.19

由前一条近似曲线得到后一条的步骤：对前一条近似曲线

---

① 参看图 23.19，在该图里，仍令 $b = \frac{1}{2}$, $a = 5$。我从 Chr. 维纳(Wiener)在 *Journ. f. Math.* 第 90 卷(1881 年)，第 221—252 页上的文章中采用这个图和其下的考察，但不采用他对魏尔斯特拉斯的结论的错误评论[维纳的文章是很有教育意义的，因为它含有不清楚的说法，而这些说法有时是联系着导数概念的。请参考魏尔斯特拉斯清晰的答复(文集第 2 卷，第 228—230 页)]。

加上一条幅度更小，而且其波长不成比例地远为更小的细密曲线，就得后一条。

现在，对于最后曲线，我们能说些什么呢？设第 m 条近似曲线已经作出，令

$$y=\sum_{0}^{\infty}b^{\nu}\cos a^{\nu}\pi x=Y_m+\sum_{m+1}^{\infty}b^{\nu}\cos a^{\nu}\pi x。$$

在总和 $\sum_{m+1}^{\infty}b^{\nu}\cos a^{\nu}\pi x$ 里，每个被加项都含有同一个因子 b^{m+1}。把它提出来，得

$$y=Y_m+b^{m+1}\sum_{0}^{\infty}b^{\nu}\cos a^{m+1+\nu}\pi x。$$

现在对右方的无尽和作出估计，以得到 y 的上界和下界，从而得到含魏尔斯特拉斯曲线在内的一条带。

令其中一切余弦的值为 $+1$，即得一个上界，令它们为 -1，即得一个下界，于是

y 的上界：$Y_m+b^{m+1}(1+b+b^2+\cdots)$，

y 的下界：$Y_m+b^{m+1}(-1-b-b^2-\cdots)$。

因为我们已明确假定正数 $b<1$，括弧里的几何级数是收敛的。于是得结果 $y=Y_m+\varepsilon b^{m+1}\cdot\frac{1}{1-b}$，其中 ε 是在 $+1$ 与 -1 之间的未知因子；在一切情况下，它包括上下界。用语言表达，就是

最后曲线的纵坐标可以对下标为 m 的近似曲线 Y_m 加上像

$$\varepsilon\cdot\frac{b^{m+1}}{1-b}\quad(-1\leqslant\varepsilon\leqslant 1)$$

形状的一项来得到。

用较多的几何语言，也可以说：

最后曲线含在一个宽度为 $2\cdot\frac{b^{m+1}}{1-b}$ 的带里，带的中线是近似曲

线 Y_m。

例如若 $b=0.1$，用厘米来量，则 $m=6$ 时，带的宽度是 $2\cdot\frac{b^{m+1}}{1-b}=2\frac{0.1^7}{0.9}=\frac{2}{9}\cdot 10^{-6}$。这样的宽度连最精密的显微镜都察觉不到。因此，若 b 适当地小，随着 m 增加，带宽非常快地减小。总之，按照这样的方法，就可以推知函数 y 的连续性。事实上有如下分析：

作为有尽多个连续函数的和，Y_m 是连续的。此外，为了得到最后曲线而加上的后面那部分，当 m 充分大时，对一切 x，它是一致地小的。但这两个事实合起来，无非是表明 y 是连续的。为了说明这个结论，我简单地把公式写下来：设

$$y(x)=Y_m+\varepsilon\frac{b^{m+1}}{1-b},$$

$$y(x')=Y_m'+\varepsilon'\frac{b^{m+1}}{1-b},$$

则

$$|y(x)-y(x')|=\left|Y_m-Y_m'+(\varepsilon-\varepsilon')\frac{b^{m+1}}{1-b}\right|,$$

$$|y(x)-y(x')|\leqslant|\underbrace{Y_m-Y_m'}_{\eta_1}|+\left|\underbrace{(\varepsilon-\varepsilon')\frac{b^{m+1}}{b-1}}_{\eta_2}\right|,$$

因而

$$|y(x)-y(x')|\leqslant\eta_1+\eta_2=\eta。$$

用语言表达，这就是

我们先取 m 如此之大，使得 $\left|(\varepsilon-\varepsilon')\cdot\frac{b^{m+1}}{b-1}\right|\leqslant 2\cdot\frac{b^{m+1}}{b-1}<\frac{\eta}{2}$，然后取 x' 如此接近 x，使 $|Y_m-Y_m'|\leqslant\frac{\eta}{2}$；这样 $|y-y'|$ 也就小于任意已给的正数 η，归结如下：

由于当 $b<1$ 时，随着 m 的增加，带宽变得任意小；另一方面，近似曲线 Y_m 是连续函数，所以当 $b<1$ 时，最后曲线是连续的。

但是，对于最后曲线，我们还可以说得更多一些。在适当选择 a 后，首先对于所谓的结和峰[①]可以给出完整的描述。

首先关于结：

当一个角是$\frac{\pi}{2}$的奇数倍时，它的余弦是零。令 $x=\frac{2g+1}{2a^m}$，其中 g 是整数。这时 $a^m\pi x=(2g+1)\frac{\pi}{2}$，故 $\cos a^m\pi x=\cos(2g+1)\frac{\pi}{2}=0$。现在再假设 a 是奇数，则由于奇数之积还是奇数，$a^{m+1}\pi x$，$a^{m+2}\pi x$，… 也具有 $(2g+1)\frac{\pi}{2}$ 的形状，于是得

$$\cos a^{m+1}\pi x=\cos a^{m+2}\pi x=\cdots=0。$$

所以，若 g 是任意整数，a 是奇数，则当 $x=\frac{2g+1}{2a^m}$ 时，不但第 m 条部分曲线有零点，在这些地方，最后曲线和第 $m-1$ 条近似曲线以及一切后面的近似曲线都有相同的纵坐标。于是在最后曲线上有一些点同时在第 $m-1$ 条近似曲线以及一切后面的近似曲线上。这些点我们称为结点。它们都在经过部分曲线的零点的纵线上。若 $2g+1$ 被 a 整除，则有关的结点已经在前面一些近似曲线上，但这种情况不需要进一步考察。

由于幂 m 以及 g 可以是任意大的整数，又因已经假定 $3a>1$，可以看出，经过结点的纵线和 x 轴的交点在 x 轴上处处稠密，即在每一个无论多小的子节内总有无尽多个结点[②]。

稍微复杂一点，但对于魏尔斯特拉斯函数的不可微性特别重要的是峰点。这个名称我们将用于近似曲线以及最后曲线的一些点。所谓近似曲线 Y_m 的峰点，是指它上面对应于 $x=\frac{g}{a^m}$ 的点，其中 g 仍然是一个

---

① 结和峰依次是 Knoten 和 Scheitel 的译名。——中译者

② 应说：在 x 轴上，每个无论多小的子节内总有无尽多个结点的投影。——中译者

整数，至于最后曲线的峰点，是指它上面一切对应于 $x=\frac{g}{a^m}$ $(m=0,1,2,\cdots)$ 的点。在这些点，第 m 条部分曲线有极大极小，因为在那里：

$$\cos a^m\pi x=\cos\pi g=(-1)^g。$$

由于 a 是奇数，因此又有

$$\cos a^{m+1}\pi x=\cos a\pi g=(-1)^g。$$

同样的结论适用于一切后面的部分曲线，于是有以下结果：近似曲线上的峰点，因而最后曲线的峰点，位于经过部分曲线的极大极小的纵线上①。

对于峰点 $x=\frac{g}{a^m}$，当 $\nu\geqslant m$ 时，魏尔斯特拉斯级数的第 ν 项的值是 $b^\nu(-1)^g$。

这样对于一个峰点，就容易给出级数的和。用 Y_{m-1} 表示第 $m-1$ 条近似曲线在那里的纵坐标，就有

$$y=Y_{m-1}+(-1)^g\cdot\frac{b^m}{1-b}。$$

把魏尔斯特拉斯曲线的一切峰点投射到 x 轴上，则由于类似关于结点的原因，这些投影在 x 轴上处处稠密。于是，到现在为止，我们的结果是：

当 $0<b<1$，a 是奇数而且大于 1 时，魏尔斯特拉斯曲线是处处连续的，而且在 x 轴上有两个确定的处处稠密的点集。作为连续函数，它的值被它在这两个点集上的值完全确定。

§ 23.10　魏尔斯特拉斯函数的不可微性

现在，下一个需要系统地考虑的问题是：

① “峰点”，这个词不完全确切；因为当 g 为奇数时，峰点对应于部分曲线的极低点。——中译者

在每个 x 处，微商的 4 个上、下界 D^+，D_+，D^-，D_- 的关系如何？我们将发现，在任何点 x，这 4 个导数没有共同的有穷值或无穷值，因而在任意 x 处，我们的函数没有有穷或无穷的导数。当然，只需证明 4 个导数中的两个，例如 D^+，D_- 有不同的值。

魏尔斯特拉斯的证明步骤是综合性的，没有对 x 作进一步的分类。为了不至于在这个问题上花太多时间，我们沿用他的方法，不过要更深入到其中的细节里去也不难。

假定我们要考察导数在 x_0 的存在性。先考虑第 m 条近似曲线，在它上面和 x_0 相邻的两个峰点中，魏尔斯特拉斯选取较靠近的一个。若 x_0 刚好在两个峰点间的正中，他选择左边的一个（图 23.20）。设 $\frac{\alpha_m}{a^m}$ 是这个峰点的横标，其中 α_m 当然是一个整数，他就得到不等式

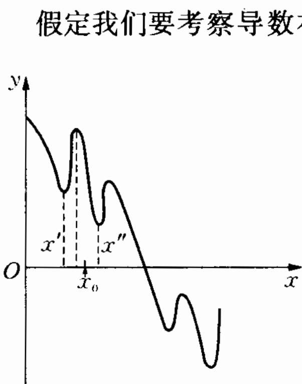


图 23.20

$$-\frac{1}{2}<x_0a^m-\alpha_m\leqslant+\frac{1}{2}。$$

这样确定的第 m 条近似曲线的峰点，本身也有两个相邻的峰点。设左边一个的横坐标是 x'，右边一个的横坐标是 x''。对于 x' 和 x''，确定最后曲线的纵标，并且把这样得到的两点 x'，y' 和 x''，y'' 同所给点 x_0，y_0 连起来。这样，我们就有最后曲线的两个差商。现在要考察的是：当 $x'-x_0$ 和 $x''-x_0$ 都趋于 0 时，它们是否接近一个共同的有穷或无穷的极限值。我们将看到，当乘积 ab 充分大时

$$\frac{y'-y_0}{x'-x_0}\text{和}\frac{y''-y_0}{x''-x_0},$$

或者趋近具有不同符号的无穷大值，或者它们都在下界 $-\infty$ 和上界 $+\infty$ 之间摆动。这样，就排除了存在有穷或无穷导数的可能性。

首先考察 $x'-x_0$ 和 $x''-x_0$。因为

$$x'=\frac{\alpha_m-1}{a^m} \text{ 和 } x''=\frac{\alpha_m+1}{a^m},$$

我们有

$$x'-x_0=\frac{\alpha_m-1-x_0a^m}{a^m}=\frac{-1-(x_0a^m-\alpha_m)}{a^m},$$

$$x''-x_0=\frac{1-(x_0a^m-\alpha_m)}{a^m},$$

或者，若按照魏尔斯特拉斯的作法，把 $x_0a^m-\alpha_m$ 简写成 x_{m+1}，则

$$x'-x_0=\frac{-1-x_{m+1}}{a^m},\ x''-x_0=\frac{1-x_{m+1}}{a^m}。$$

由于 $|x_{m+1}|\leqslant\frac{1}{2}$，又 $a>1$，因此当 m 增加时，两个差都趋于零。最后曲线的峰点集合对应于 x 轴上一个处处稠密的点集，这本来是可以立刻看得出的。

把 y' 和 y'' 写成

$$y'=\sum_{\nu=0}^{m-1} b^\nu\cos a^\nu\pi x'-(-1)^{\alpha_m}b^m\sum_{\nu=0}^{\infty} b^\nu,$$

$$y''=\sum_{\nu=0}^{m-1} b^\nu\cos a^\nu\pi x''-(-1)^{\alpha_m}b^m\sum_{\nu=0}^{\infty} b^\nu。$$

这时，若把 $y'-y_0$ 分成两部分，则差商

$$\frac{y'-y_0}{x'-x_0}=\sum_{\nu=0}^{m-1} b^\nu\frac{\cos a^\nu\pi x'-\cos a^\nu\pi x_0}{x'-x_0}$$

$$+\sum_{\nu=0}^{\infty} b^{m+\nu}(-1)^{\alpha_m+1}\frac{1+\cos a^\nu\pi x_{m+1}}{x'-x_0}。$$

把 $y'-y_0$ 分为两部分的想法是，把第 $m-1$ 条近似曲线 $\sum_0^{m-1} b^\nu\cos a^\nu\pi x$

和“余曲线”$\sum_{m}^{\infty} b^{\nu}\cos a^{\nu}\pi x$ 的差商分开。现在，对两个差商分别进行估计。

在第 $m-1$ 条近似曲线的差商里，把分子中的差化为积，就得

$$\sum_{\nu=0}^{m-1} b^{\nu} \frac{\cos a^{\nu}\pi x' - \cos a^{\nu}\pi x_0}{x' - x_0}$$

$$= \sum_{\nu=0}^{m-1} b^{\nu} \frac{2\sin a^{\nu}\pi \dfrac{x_0 + x'}{2} \cdot \sin a^{\nu}\pi \dfrac{x_0 - x'}{2}}{x' - x_0}。$$

以 $\frac{1}{2}a^{\nu}\pi$ 乘第 ν 项的分子和分母，就得

$$\sum_{\nu=0}^{m-1} b^{\nu} \frac{\cos a^{\nu}\pi x' - \cos a^{\nu}\pi x_0}{x' - x_0}$$

$$= \sum_{\nu=0}^{m-1} -a^{\nu}b^{\nu}\pi \frac{\sin a^{\nu}\pi \dfrac{x_0 - x'}{2}}{a^{\nu}\pi \dfrac{x_0 - x'}{2}} \cdot \sin a^{\nu}\pi \frac{x_0 + x'}{2}。$$

由 $\left|\dfrac{\sin a^{\nu}\pi \dfrac{x_0 - x'}{2}}{a^{\nu}\pi \dfrac{x_0 - x'}{2}}\right| \leqslant 1$ 和 $\left|\sin a^{\nu}\pi \dfrac{x_0 + x'}{2}\right| \leqslant 1$ 可知，第 $m-1$ 条近似曲线的差商的绝对值小于或等于 $\pi\sum_{0}^{m-1} a^{\nu}b^{\nu} = \pi\dfrac{a^m b^m - 1}{ab - 1}$，而且肯定大不到 $\pi\dfrac{a^m b^m}{ab - 1}$（已假定 $ab > 1$）。因此可以令第 $m-1$ 条近似曲线的差商等于

$$\varepsilon\pi \frac{a^m b^m}{ab - 1} \quad (-1 < \varepsilon < +1),$$

至于余曲线的差商，则

$$\sum_{\nu=0}^{\infty} b^{m+\nu}(-1)^{\alpha_m+1} \frac{1+\cos a^{\nu}\pi x_{m+1}}{x'-x_0}$$

$$=(-1)^{\alpha_m}a^m b^m \sum_0^{\infty} b^{\nu} \frac{1+\cos a^{\nu}\pi x_{m+1}}{x_{m+1}+1}。$$

这里面不尽级数第一项是

$$\frac{1+\cos\pi x_{m+1}}{x_{m+1}+1}。$$

由于

$$-\frac{1}{2}<x_{m+1}\leqslant+\frac{1}{2},$$

$\cos\pi x_{m+1}\geqslant 0$。分母 $x_{m+1}+1$ 则在$\frac{1}{2}$ 和$\frac{3}{2}$ 之间摆动。因此

$$\frac{1+\cos\pi x_{m+1}}{x_{m+1}+1}\geqslant\frac{2}{3}。$$

级数后面各项或正或等于零,所以级数的和也大于等于$\frac{2}{3}$。

设 η'为大于等于 1 的一个正数,并令 $\frac{(-1)^{\alpha_m}}{\eta'}\cdot\varepsilon=\varepsilon'$(这样,$\varepsilon'$ 像 ε 那样也在 -1 与 $+1$ 之间),就有

$$\frac{y'-y_0}{x'-x_0}=(-1)^{\alpha_m}a^m b^n \eta'\left(\frac{2}{3}+\varepsilon'\frac{\pi}{ab-1}\right)。$$

于是得到魏尔斯特拉斯函数左方差商的一个估计,类似地,对右方差商有公式

$$\frac{y''-y_0}{x''-x_0}=(-1)^{\alpha_m+1}+a^m b^m \eta''\left(\frac{2}{3}+\varepsilon''\frac{\pi}{ab-1}\right),$$

不过这里多了一个因子(-1),因为 $x''-x_0$ 是正的而 $x'-x_0$ 是负的。

我们希望能使差商中余曲线提供的那部分的绝对值大于第 $m-1$条近似曲线提供的那部分的绝对值。通过定性考虑,我们可得

结论：

必须使得加到第 $m-1$ 条近似曲线上的那些波状线尽可能地陡。但陡度与 ab 有关，因此必须令 ab 充分地大。

定量地可实现如下：

我们选取不利的款：ε'（以及 ε''）等于 -1，这时必有 $\frac{2}{3}>\frac{\pi}{ab-1}$，即

$$ab>1+\frac{3\pi}{2}。$$

现在假定这个条件已得到满足。这时 $\eta'\left(\frac{2}{3}+\varepsilon'\frac{\pi}{ab-1}\right)$肯定是一个依赖于 m 的正数 p_m'。同样有

$$\eta''\left(\frac{2}{3}+\varepsilon''\frac{\pi}{ab-1}\right)=p_m''。$$

于是得

$$\left.\begin{aligned}p_m'&=\eta'\cdot\frac{2}{3}+\eta'\cdot\varepsilon'\frac{\pi}{ab-1}\\p_m''&=\eta''\cdot\frac{2}{3}+\eta''\cdot\varepsilon''\frac{\pi}{ab-1}\end{aligned}\right\}\geqslant\left(\frac{2}{3}-\frac{\pi}{ab-1}\right)。$$

若令 $\frac{2}{3}-\frac{\pi}{ab-1}=q$，则

$$\left.\begin{aligned}p_m'\\p_m''\end{aligned}\right\}\geqslant q,$$

其中 q 的值与 m 无关。

于是关于差商，有以下两种关系之一：

(1) $\dfrac{y'-y_0}{x'-x_0}\geqslant(-1)^{\alpha_m}a^mb^mq$，

$$\frac{y''-y_0}{x''-x_0}\leqslant(-1)^{\alpha_m+1}a^m b^m q;$$

(2) $$\frac{y'-y_0}{x'-x_0}\leqslant(-1)^{\alpha_m}a^m b^m q,$$

$$\frac{y''-y_0}{x''-x_0}\geqslant(-1)^{\alpha_m+1}a^m b^m q,$$

依次按照 $(-1)^{\alpha_m}=+1$ 或 $(-1)^{\alpha_m}=-1$ 而定。

这些就是魏尔斯特拉斯证明导数不存在性所依据的最后公式。证明的具体步骤如下：在差商中，令 m 越来越大，使 x' 从左方，x'' 从右方无限制地接近 x_0。若假定有穷导数存在，则两个差商都必须无限制地趋近同一个有穷数。若假定无穷导数存在，则当 $|x'-x_0'|$ 和 $|x''-x_0|$ 充分小时，两个差商最后保持以相同符号向无穷大增长。但令 m 加大时，从两个差商的最后公式可知，要区别 3 款：

(1) 一切 α_m（除有尽多个外）是偶数；这时

$$\frac{y'-y_0}{x'-x_0}\text{ 趋于 }+\infty,\ \frac{y''-y_0}{x''-x_0}\text{ 趋于 }-\infty。$$

(2) 一切 α_m（除有尽多个外）是奇数；这时

$$\frac{y'-y_0}{x'-x_0}\text{ 趋于 }-\infty,\ \frac{y''-y_0}{x''-x_0}\text{ 趋于 }+\infty。$$

(3) 若既非第(1)款又非第(2)款，则可以假设，当 m 增大时，$(-1)^{\alpha_m}$ 有交错符号，无损于普遍性。但这时 $\dfrac{y'-y_0}{x'-x_0}$ 和 $\dfrac{y''-y_0}{x''-x_0}$ 分别都是在以 $-\infty$ 为最大下界，以 $+\infty$ 为最小上界之间摆动。所以在任何一款，魏尔斯特拉斯函数在 x_0 都没有有穷或无穷导数，而 x_0 则完全是任意的。

我们用一个图来说明问题的实质。取第 m 条近似曲线上一个下峰①（α_m 为奇数）所对应的 x_0（图 23.21），它也是后面的一切近似

① 下峰即极小点，称为"倒峰"可能更恰当。——中译者

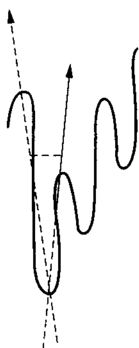


图 23.21

曲线的下峰。这是因为一切 $\alpha_{m+\nu}(\nu=0,1,2,\cdots)$ 都是奇数。于是当 m 增长时,它和相邻的峰点的高度差必然缩小,它们和 x_0 的水平距离也缩小,但由于 ab 选择得那么大,高度差缩小的速度比水平距离缩小的慢得多。其结果是,当 m 增长时,对应于差商的两条弦无限制地变得更陡,因而它们的斜率以相反符号变得无穷大。上面是假设了第 m 条部分曲线的一个下峰同时是第 $m+1$ 条,第 $m+2$ 条,…部分曲线的下峰。若 $\alpha_{m+\nu}(\nu=0,1,2,\cdots)$ 是偶数,x_0 当然对应于有关部分曲线的上峰,而和它相邻的峰则是下峰,于是图像倒过来了,但假定导数存在仍然引出矛盾①。

§ 23.11 "合理"函数②

关于魏尔斯特拉斯函数的讨论表明,假定了函数连续还不能保证导数存在。若要求连续函数有一阶、二阶以及更高阶导数,就必须把这些条件明确地提出。

回顾本章的论述,有如下内容:我们从经验曲线自然地看到某些性质:

(1) 连续性;

① [读者会猜测,对于魏尔斯特拉斯函数的不可微性,条件 $ab>1$ 已经够了,它可以取代所采用的更强的条件 $ab>1+\frac{3\pi}{2}$,而且条件"a 是奇数"与问题的本质无涉。实际上,G·H·哈代(G. H. Hardy)(*Transactions of the American Mathematical Söciety*,第 17 卷,1916 年,第 301—325 页)已证明了,如果把不可微性理解为不存在有穷导数,则条件 $0<b<1, ab\geqslant 1$ 就够了。若不限于有穷导数,要保证不可微性,上述条件就不够了。——这个事实,魏尔斯特拉斯是知道的。从第 46 页所提到的他给柯尼希斯伯格的信中可以看出。我们上面在 $0<b<1$ 外,还要求 a 是奇数,$ab>1+\frac{3\pi}{2}$。这也有优点,可以使用初等方法来论证。]

② vernünftige Funktionen. ——中译者

(2) 在一个有穷节内有一个最大值和一个最小值，以及有尽多个极大值和极小值；

(3) 存在着方向；

(4) 存在着曲率。

若要求精确数学里的函数也有类似表现，就必须明确假定它们有下列性质：

(1) 在一个闭节里的连续性；

(2) 在一个闭节里有有尽多个极大值和极小值；

(3) 存在着一阶导数；

(4) 存在着二阶导数，等等。

这样就从全部函数中筛选出一类完全确定的函数，但它们比解析函数要普遍些，因为并没有要求它们有任意高阶的导数，更不要说可以用泰勒级数表示了。我沿用雅可比(Jacobi)的名词，把它们叫做"合理函数"①。

引进这个名词，就可以说：

经验曲线通常定性地具有的性质，在精确数学中我们称为合理函数的那类函数 $y=f(x)$上重现了。

这里还没有涉及两者定量方面的一致性问题，这个问题将在下一章讨论。

① 下文将要提到的正则函数与正则曲线，意义与此相同。

第二十四章
函数的近似表示

§24.1　用合理函数近似表示经验曲线

设已给一条经验曲线。问题是：能否在精确数学中确定一个函数 $y=f(x)$，使得对每个 x，它在纵坐标、方向和曲率等方面都和经验曲线充分接近？为此目的，我们将不使用复杂的函数而只用简单的、用解析式代表的函数。因此，紧接着上述问题就可提出另一个问题：

在多大程度上，可以在路线整体、方向、曲率等方面，用一个简单的、由解析式界定的函数来近似地表示一条经验曲线？

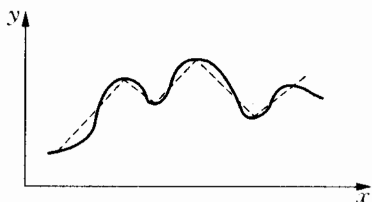


图 24.1

首先我们通过一个任意的具体例子来说明，总可以用合理函数来任意近似地代表每一条经验曲线。假定要求用函数 $y=f(x)$ 来充分准确地表示图 24.1 所规定的纵标和横标的关系。我们取一个边数充分大的内接多边形，使它的各边和其对应的曲线段明显地吻合。在各具体的款，如何取多边形，或者说，如何选取其各边的长及其端点，在这里提不出一般办法，但总可以用多边形来代替每条经验曲线，使它达到明显的准确度①，而这个多边形就代表了一个公式 $y=f(x)$，这个公式就

① 这是从实际经验得到的方法，而且它在数学思想中占有重要地位。

可以看做曲线所代表的规律的近似表示。

但这样我们并没有得到一个处处可微函数，更不要说该函数的导数处处和曲线的"方向"吻合了；它和经验曲线在误差要求限度内的吻合只限于纵坐标。

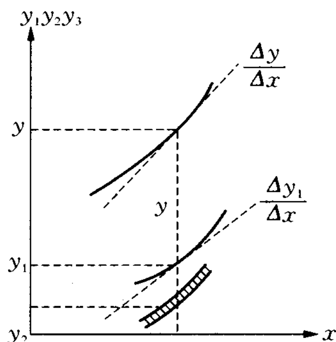


图 24.2

为了解决关于方向的问题，我们采取下述步骤。设 y 是所给经验曲线的纵坐标，在每一处 x，取确定曲线方向的差商（参看第 28 页）$\frac{\Delta y}{\Delta x}$，取$\frac{\Delta y}{\Delta x}$的值 y_1 作为一条新曲线 $y_1=f_1(x)$的纵标。这条新曲线就叫做"第一导出曲线"①（图 24.2）。

由于经验曲线只是用一个"带"表示而不是准确地给定的，因此方向$\frac{\Delta y}{\Delta x}$可以在较宽的幅度中游动，"第一导出曲线"有较明显的宽度；但总要受到一种制约，即（在允许的准确限度内）$\int f_1(x)\mathrm{d}x=f(x)$。

类似地可以作"第二导出曲线"，即以第一导出曲线的方向$\frac{\Delta y_1}{\Delta x}$作为纵坐标所得曲线。它当然有更大的不准确度，等等。

在完成以上步骤之后，我们用多边形代替曲线 y_2，以得到一个函数 $f_2(x)$，对它积分两次，就得到一个函数 $f(x)$，这函数在纵坐标、方向和曲率等方面都以所期望的准确度代表所给曲线。用公式简短地表示：

$$y_2=f_2(x),\quad y_1=\int f_2(x)\mathrm{d}x,\quad y=\int y_1\mathrm{d}x=f(x)。$$

① abgeleitete Kurve. ——中译者

归纳起来：

为了确定一个合理函数，使它在所给准确度范围内，不但代表已给经验曲线的纵标，也代表其方向和曲率，我们在实际可能限度内，作所给曲线的第一和第二导出曲线，然后用多边形代替第二导出曲线以得到函数 $f_2(x)$，把 $f_2(x)$ 积分两次，就得到所求的近似函数。

当然，这是获得具有所期望性质的合理函数的一种方法；每个数学工作者都能立刻想出别的方法。

在具体工作中，如何运用这样一种方法，只能根据情况作出实际判断。在这里，我们绝不仅仅是在纯粹数学的严格逻辑领域中处理问题，而是在这样一个领域中处理问题：在这里，除数学的纯粹逻辑推理外，某些应用性、目的性和可行性的考虑也起着重要作用。由于我们既不能也不想深入到具体的应用领域中去（在那里，如上面所看到的，总有那么大的不确定性），在进一步讨论用简单的解析式作近似表示时，我们不再设想已给的是一条经验曲线而是一个“合理函数”$y=f(x)$。这个函数可以是一次或二次（或更多次）可微的。我们的问题是：这样一个合理函数在多大程度上可以用一个简单的解析式（多项式、三角函数和等等）来逼近？

这样，我们就把关于近似表示问题的探讨完全转入纯数学领域；一切假设都是严格地提出的，从而一切论点都是严格地陈述的。至于“应用”意义，就要看我们的论述中哪些地方能起实际作用了。

§24.2 用简单解析式近似表示合理函数

用以作近似表示的简单解析式首先是

(a) 多项式

$$y = A + Bx + Cx^2 + \cdots + Kx^n,$$

(b) 有尽三角级数（三角函数多项式或和）

$$y = a_0 + a_1\cos x + a_2\cos 2x + \cdots + a_\nu\cos nx$$
$$+ b_1\sin x + b_2\sin 2x + \cdots + b_\nu\sin nx。$$

我们要考察的是,一个合理函数在多大程度上可以用这些简单的函数表示,而且特殊地,导数的接近情况又如何。

在课本里,用有尽级数近似表示的问题常常让位于无尽级数(泰勒级数、傅里叶级数等)的准确表示。但后者完全是另外一个问题。它本身的确是很重要的,但在应用中它从来不起作用。因为在应用中所要解决的,自然是在多大程度上以及在什么意义上可以用有尽级数去逼近。许多课本的这种片面性只能理解为作者不是从实践出发而是从理论基础着眼的。

在近似表示中可以有两种想法:

(a) 多项式以及有尽三角级数只是在某些地方与函数和它的导数吻合;

(b) 选取级数使所得到的结果在整体上按照例如最小二乘法的要求,令其误差平方的和为极小①。

§24.3　拉格朗日插值公式

我从多项式开始,它们只在若干已给处和所给函数吻合。

设 $y=f(x)$,另取 n 个 x 值 $x=\alpha, \beta, \cdots, \nu$,在这些处,希望所求多项式具有所给函数的纵坐标。

我们将采用熟知的拉格朗日插值公式,它给出满足要求的最低次多项式,只含有 n 个常数:

$$Y = f(\alpha)\frac{(x-\beta)(x-\gamma)\cdots(x-\nu)}{(\alpha-\beta)(\alpha-\nu)\cdots(\alpha-\nu)}$$

---

① [上面所提的问题,在第一卷(第 205—215,241—255 页)(中译本第 217—229,256—273 页)已经讨论了,为了下面的目的,不需要复习那里的论述。]

$$+ f(\beta)\frac{(x-\alpha)(x-\gamma)\cdots(x-\nu)}{(\beta-\alpha)(\beta-\gamma)\cdots(\beta-\nu)}$$

$$+\cdots+ f(\nu)\frac{(x-\alpha)(x-\beta)\cdots}{(\nu-\alpha)(\nu-\beta)\cdots}。$$

它是一个 $n-1$ 次多项式,在所给各处,它和函数 $y=f(x)$ 的确有相同的纵坐标。例如,当 $x=\beta$ 时,$Y=f(\beta)$。

现在要考察,在所给 n 处以外,拉格朗日插值多项式和我们的函数接近到什么程度?

为了作出判断,我们用 $\Theta(x)$ 表示多项式 Y,并用 $R(x)$ 表示余项 $y-Y$,则

$$y=\Theta(x)+R(x)。$$

由于当 $x=\alpha, \beta, \cdots$ 时,$R(x)$ 等于零,可以把因子 $\varphi(x)=(x-\alpha)(x-\beta)\cdots(x-\nu)\cdots$ 提出来,因此得

$$y=\Theta(x)+\varphi(x)\cdot r(x)。$$

同时拉格朗日公式就可以写成

$$\Theta(x)=\frac{f(\alpha)}{\varphi'(\alpha)}\frac{\varphi(x)}{x-\alpha}+\cdots+\frac{f(\nu)}{\varphi'(\nu)}\frac{\varphi(x)}{x-\nu}。$$

若要拉格朗日公式在已给节里能用,$|r(x)|$ 在这个节里必须保持充分地小,$\varphi(x)$ 在每个闭节里是有界的。于是我们有以下的中心问题:

能否把 $r(x)$ 限制在一定幅度里,使得 $\Theta(x)$ 可用作 $f(x)$ 的一个近似式?

我们这里所讨论的内容通常称为内插法,这个名词的来源是,我们本来设想,把对 $f(x)$ 的近似表示只限于在节 $\alpha, \beta, \cdots, \nu$ 里的 x 值。可是我们的问题也涉及在这个节外的 x(外推法)①。

---

① [在第一卷第 247 页(中译本第 263 页)上曾指出应当把“内插法”换成把“外推法”包含在内的“近似”一词,可是最近内插已经无形中把外推包括在内了。]

从这个一般的讨论,我们要进入两个或更多的点 α, β, …重合的特殊情况,即在某些地方,一阶或更高阶的导数已经给定的情况①。

若除了 $f(\alpha)$, $f(\beta)$,…以外,还给定了 $f'(\alpha)$, $f'(\beta)$,…,就有如下问题:

如何作出一个多项式,它在 α, β, γ,…不但有所要求的纵坐标,还有所给的导数(密切插值)?

这个问题我们可以直接处理,也可以由拉格朗日公式通过极限过程解决。下一个问题是,这样所得到的多项式在多大程度上可以用来近似表示 $f(x)$和它的导数?

特殊地,令一切点 α, β, …, ν 重合到一个点 a,就得到泰勒公式

$$f(x) = f(a) + \frac{f'(a)}{1}(x-a) + \frac{f''(a)}{2!}(x-a)^2$$

$$+ \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}(x-a)^{n-1} + r(x) \cdot (x-a)^n。$$

从拉格朗日公式不难推得这个结果,在这里我们不加说明②。

现在对拉格朗日公式和它的特款,我们特别感兴趣的是,如上边已经提出的,余项的估计。

作为余项估计的基础,我们利用罗尔(Rolle)定理,它是微分学里中值定理的特款。罗尔定理说:设 $F(z)$是在闭节 $a \leqslant z \leqslant b$ 连续的函数,它在节内部的每一点有导数,此外,设 $F(a)=F(b)=0$。则导数 $F'(z)$在节的内部至少有一个零点(图 24.3)。

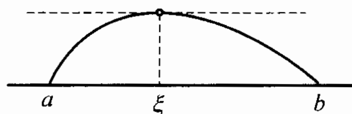


图 24.3

这个定理的证明是容易的。

① [进一步内容见第一卷第 247—252 页(中译本第 263—268 页)。]
② [详尽讨论见第一卷第 247(中译本第 263)等页。]

按照魏尔斯特拉斯定理，每一个在闭节里连续的函数在节的内部或在节的一个端点有最大值。我们不考虑 $F(z)=$常数的平凡款，并且先假定 $F(z)$在节 $a\cdots b$ 内部的一些地方有正值。这时 $F(z)$必在节的内部某处 ξ 有最大值。根据导数的唯一性，立刻就得 $F'(\xi)=0$。

若 $F(z)$在节的内部只有负值，我们就考虑$-F(z)$的最大值，从而得到相同结论。

如你们所看到的，这个定理的简单基础是魏尔斯特拉斯定理和 $F(z)$的一次可微性。

下面我们把定理推广到 $F(z)$在 3 点 a, b, c 等于 0 的款。应用我们的定理两次，可知在节 $a\cdots c$ 里 $F''(z)$至少有一次等于零。这样就得以下的定理，我把它写成适合后面应用的形式：

设 $F(z)$有 k 个零点，再设 $F(z)$在一个含有这些零点的一个闭节里连续，而且充分多次可微，则在该节里，$F(z)$的 $k-1$ 阶导数至少有一个零点。

我们按下述方法利用这个定理来估计拉格朗日插入公式中的余项。

考虑函数

$$F(z)=f(z)-\Theta(z)-r(x)\cdot\varphi(z),$$

其中$\Theta(z)$[$(n-1)$次]表示不带余项的拉格朗日多项式，$\varphi(z)$（n次）是第 66 页所引用的因子，而 x 是任意选定的自变数 z 的一个固定值。现在，我们知道函数 $F(z)$有一系列的零点。首先，当 z 等于 $\alpha, \beta, \cdots$ 时，$f(z)-\Theta(z)$ 和 $\varphi(z)$ 都等于零，因而 $F(z)=0$。此外，x 也是零点，因为当 $z=x$ 时，根据定义 $f(x)=\Theta(x)+\gamma(x)\cdot\varphi(x)$；所以 $F(x)=0$。

于是我们可以把一般形式的罗尔定理应用于 $F(z)$。在这里，我们令 $k=n+1$。根据该定理，$F^{(n)}(z)$在节 $\alpha, \cdots, \nu, x$ 内部至少有一个零点 ξ。把这个 n 阶导数算出来，得

$$F^{(n)}(z) = f^{(n)}(z) - r(x) \cdot n!。$$

[$\Theta(z)$作为一个 $n-1$ 阶多项式,它的 n 阶导数是零,$r(x)$是常数,$\varphi(z)=z^n+\cdots$的 n 阶导数是 $u!$]。所以对于上述的 ξ

$$f^{(n)}(\xi) - r(x)n! = 0,$$

因而

$$r(x) = \frac{f^{(n)}(\xi)}{n!}。$$

把这个通过如此简单但却富有意义的方法所得到的 $r(x)$的值代进带余项的拉格朗日公式,就得

$$f(x) = \Theta(x) + \varphi(x)\,\frac{f^{(n)}(\xi)}{n!},$$

其中的 ξ 除了知道它在那个节 $\alpha, \beta, \cdots, \nu, x$ 内某处之外,是未知量。

于是我们有以下结果:

只要 $\varphi(x) \cdot f^{(n)}(\xi): n!$ 对于节里的一切 ξ 是一个充分小的值,拉格朗日公式就是可以用的,即,$f(x)$可以用 $\Theta(x)$近似地表示无论 x 在 $\alpha\cdots\nu$ 之间或者之外,这话都适用。这个带余项的公式对于内插或外插同样适用。

§ 24.4　泰勒定理和泰勒级数

特殊地,若要同泰勒公式联系起来,就有

$$f(x) = f(a) + \frac{f'(a)}{1}(x-a) + \cdots$$
$$+ \frac{f^{(n-1)}(a)}{(n-1)!}(x-a)^{n-1} + \frac{f^{(n)}(\xi)}{n!}(x-a)^n。$$

这里的余项和通常在微积分学中论述泰勒级数时所得的拉格朗日余项形式相同。

泰勒公式只涉及外插，因为在这里，所有 $\alpha, \beta, \cdots, \nu$ 都重合到同一点 a。

我现在通过一系列的例子来说明上面所讨论内容。

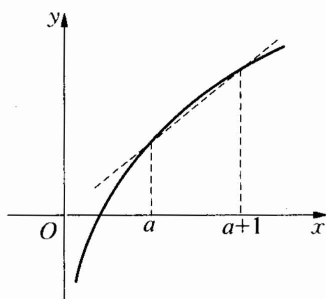


图 24.4

首先我们考察，在使用对数表中拉格朗日公式的应用。在对数表中，我们先找到两个数 a 和 $a+1$。问题是：我们能否利用直线来进行内插，即能否把纵坐标 $\log a$ 和 $\log(a+1)$ 之间的那段对数曲线用它的弦来代替(图 24.4)?

令 $f(x)=\log x$①，则当 $n=2$ 时，拉格朗日公式是

$$\log x=\log a+(x-a)\{\log(a+1)-\log a\}+R。$$

为了估计余项

$$R=\frac{f''(\xi)}{2}(x-a)(x-a-1),$$

我们有 $f'(x)=\frac{\mathrm{d}\log x}{\mathrm{d}x}=\frac{M}{x}$ 和 $f''(x)=-\frac{M}{x^2}$，其中 M 是常用对数的模②。乘积 $(x-\alpha)(x-\alpha-1)$ 的曲线图像是一条抛物线。在节 $a\leqslant x\leqslant a+1$ 里，当 $x=\frac{a+(a+1)}{2}=a+\frac{1}{2}$ 时，它的纵坐标的绝对值最大，这时纵坐标等于 $-\frac{1}{4}$(图 24.5)。把它代入上面的公式，就

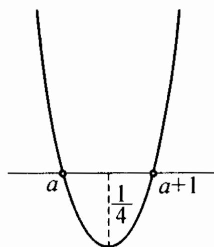


图 24.5

① $\log x$ 是常用对数 $\log_{10}x$。——中译者

② $M=\log_{10}e$。——中译者

可看出略去余项所产生的误差的绝对值$|r(x)|$的上界是$\frac{M}{8\xi^2}$。

所以，当我们用比例方法来计算一个对数值时，所得到的值是小了些，但最多只差$\frac{M}{8\xi^2}$。在这里，ξ是a和$a+1$之间的任何值。因此，误差的绝对值肯定小于$\frac{M}{8a^2}$。

我们再通过一个数值的例子来说明误差的大小。假定我们用的是七位对数表，在那里，对数的值已经给出了五位数。这样，a以及ξ都是五位数。对于常用对数的模$M=0.434\ 29\cdots$我们用一个较大的数0.5替代，当选取a为最小的五位数时(这时误差最大)，就得到误差的上界是$\frac{0.5}{8\cdot 10\ 000^2}=\frac{1}{16\cdot 10^8}$。这比用五位对数表算七位对数值所产生的误差还要小[①]。

作为第二个例子，我们对泰勒定理

$$f(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\cdots$$
$$+\frac{f^{(n-1)}(a)}{(n-1)!}(x-a)^{n-1}+\frac{f^{(n)}(\xi)}{n!}(x-a)^n$$

作一次说明。当$f(x)$不是无限制地可微时，这个公式仍然有良好的意义。另一方面，在所有课本里，通常假定了$f(x)$是无限制可微的。这时，当n增大时，余项的绝对值无限制地减小，$f(x)$就等于这样所得到的无尽级数。在众多情况下，当$|x-a|$充分小时，就是如此；这时，我们就有一个“收敛节”。若级数对于一切有穷x值收敛，$f(x)$就称为一个整函数。

---

① [克莱因在最后一次(1911年)关于微积分的讲演中详尽地讨论了，若作插入法时，用的不是$\log a$和$\log(a+1)$的准确值而是近似值，如何求得误差。在这个可惜未发表的讲演中，克莱因以优美的方式运用了函数带的思想来引导我们的直观。对此，洛莱(W. Lorey)在*Zeitschr. f. Math. u. naturw. Unterricht* 第43卷(1912年)，第544—556页有报道。]

当无尽级数有收敛时,考察泰勒级数的部分和所对应的近似曲线的情况是有意义的。由于这在第一卷里已经用图形详细说明过了,因此这里就无需再次回到这个问题①。

§ 24.5 用拉格朗日多项式近似表示积分和导函数

现在我们进一步讨论以下问题:带着余项的拉格朗日公式在多大程度上可以用来得到一个函数 $f(x)$的积分和它的导函数的近似表示?

先考虑积分的近似表示。

在多数课本里,这是最重要的,因为拉格朗日公式普遍用来作一条曲线的面积的数值运算(所谓的机械求积法或数值积分法)。可是误差的估计往往被省略了或者没有进行像公式实际运用中所需要的那样深入地讨论②。这里我们只对 4 个重要而简单的款作简短概括,但不深入到具体的运算。

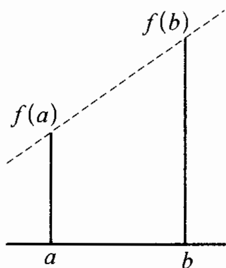


图 24.6

假定要计算积分 $\int_a^b f(x)\mathrm{d}x$。

(a) 假定在 a 和 b,函数 $f(x)$的值已经给定,在纵坐标 $f(a)$和 $f(b)$之间,用弦去替代那一段曲线(图 24.6)。在拉格朗日公式里令 $n=2$,就得

$$\int_a^b f(x)\mathrm{d}x = \frac{f(a)+f(b)}{2}(b-a) - \frac{f''(\xi)(b-a)^3}{12},$$

① 第一卷第 241—245 页(中译本第 256—260 页)。

② [参看诸如龙格和康尼格:*Numerisches Rechnen*(《数值计算》,柏林,1924 年);J. F. Steffensen:*Interpoiationslaere*(斯特芬森:《插入理论》,哥本哈根,1925 年)(英文版:*Interpolation*(《插入法》),Baltimore,1927 年;德文版在准备中。)]

即这个积分等于梯形面积加上一个余项，其中 ξ 是在节 $a\cdots b$ 内的一个未知量。

(b) 另一个用直线来插入的方法，是用在 $x=\frac{a+b}{2}$ 处的切线来代替曲线。这时就得

$$\int_a^b f(x)\mathrm{d}x = f\left(\frac{a+b}{2}\right)(b-a)+\frac{f''(\xi)(b-a)^3}{24},$$

这里的误差和上一款的误差相比，减少一半，而且有相反的符号。

值得注意的是，在最后的公式里，斜率 $f'\left(\frac{a+b}{2}\right)$ 没有出现。事实上，对于用以作为近似面积的那个梯形上，那条经过点 $x=\frac{a+b}{2}$，$y=f\left(\frac{a+b}{2}\right)$ 的边的方向，不影响梯形面积(图 24.7)。

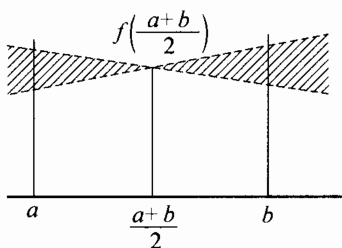


图 24.7

(c) 写下当 $n=4$ 时的拉格朗日公式就是用一条三次抛物线来近似表示曲线 $y=f(x)$，可以用 a，$f(a)$ 和 b，$f(b)$ 两点以及曲线在这两点的方向来确定那条抛物线，于是得

$$\int_a^b f(x)\mathrm{d}x = \frac{f(a)+f(b)}{2}(b-a)$$
$$-\frac{f'(b)-f'(a)}{12}(b-a)^2+\frac{f^{\mathrm{IV}}(\xi)(b-a)^5}{720}。$$

(所谓欧拉的"和公式"的最简单款)

(d) 确定三次抛物线的另一种方法是用纵坐标 $f(a)$ 和 $f(b)$ 的

终点和纵坐标 $f\left(\frac{a+b}{2}\right)$ 的终点以及那里已知的方向 $f'\left(\frac{a+b}{2}\right)$（图

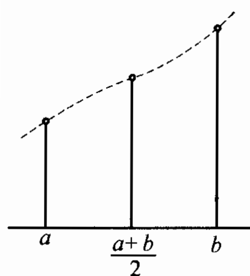


图 24.8

24.8[①]）。这时 $f'\left(\frac{a+b}{2}\right)$ 在最后公式里还是不出现。结果是

$$\int_a^b f(x)\mathrm{d}x = \frac{f(a)+4f\left(\frac{a+b}{2}\right)+f(b)}{6}(b-a) - \frac{f^{\mathrm{IV}}(\xi)(b-a)^5}{2\,880}。$$

最后这一行含有利用所谓辛普森定则所得到的面积公式[②]。

为了考察在多大程度上带余项的拉格朗日公式近似地表示 $f(x)$ 的导函数问题，我们谈论一个结果。

我们知道函数

$$f(x)-\Theta(x),$$

有 n 个零点 $\alpha, \beta, \cdots, \nu$。为了比较 $f'(x)$ 和 $\Theta'(x)$，我们考察函数 $f'(x)-\Theta'(x)$ 的零点。

由中值定理，我们立刻可知，它在 $\alpha, \cdots, \nu$ 之间至少有 $n-1$ 个根 $\alpha', \beta', \cdots, \mu'$。若 $\psi(x)$ 表示乘积 $(x-\alpha')\cdots(x-\mu')$，可以令

$$f'(x)=\Theta'(x)+s(x)\cdot\psi(x),$$

其中 $s(x)$ 可以像 $\gamma(x)$ 那样计算（参看第 67—69 页）。容易推得

$$s(x)=\frac{f^{(n)}(\xi)}{(n-1)!},$$

---

① 此处为编辑所加。

② [关于机械求积、插入法理论以及数值计算等的详尽著作，除上面已提出的龙格和柯尼希以及斯蒂芬森的书外，还可以特别举出惠特克（E. T. Whittaker）和罗宾森（G. Robinson）的 *The calculus of observations*（《观测的数学》，伦敦，1924 年）]。

因而有以下结果。

我们可以令

$$f(x)=\Theta(x)+\frac{f^{(n)}(\xi)}{n!}\cdot\psi(x),$$

以得到良好结果；对于导函数，有近似公式

$$f'(x)=\Theta'(x)+\frac{f^{(n)}(\xi)}{(n-1)!}\psi(x),$$

其效果也同样好，其中 $\psi(x)$ 是新因子 $(x-\alpha')$，…，$(x-\mu')$ 之积而 ξ 是节 α'，…，μ'，x 里的一个未知量。

对具体的款，我们当然要设法求出函数 $\psi(x)$。当 $\psi(x)$ 不能具体确定时，有时可以对它的值作出估计①。

§ 24.6　关于解析函数及其在阐释自然中的作用

在此，我想插进关于解析函数的一些一般性说明，并指出它在多大程度上和当前讨论的内容相联系或者相反地，不相联系。

我们从泰勒公式出发

$$f(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\cdots$$
$$+\frac{f^{(n-1)}(a)}{(n-1)!}(x-a)^{n-1}+\frac{f^{(n)}(\xi)}{n!}(x-a)^{n}。$$

在这里，我们明确指出，带余项的泰勒公式不假定函数 $f(x)$ 是无限制地可微的，而只假定它是 n 次可微的。为明确起见，我还要加一句：在应用中，只有在所考虑的节里，这个公式的余项可以略去时，

① ［关于数学上十分有趣的数值微导的问题，可参看第 74 页所提到的斯蒂芬森的书。］

公式才适用。

什么时候 $f(x)$ 称为解析函数?

它需要满足两个条件:

显然泰勒公式必须形式上能无限制地延伸,即任意高阶导数必须存在,而且当 n 增大时,余项最后必须变得任意小。

当这些条件得到满足时,我们才能在某个节内令

$$f(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\frac{f''(a)}{2!}(x-a)^2+\cdots,$$

它有无尽多项。

应当把含有尽多项以及余项的公式称为泰勒定理,而把含无尽多项的公式称为泰勒级数。于是可以说:

若 $f(x)$ 在含 $x=a$ 的一个节里可用泰勒级数表示,它就称为在 $x=a$ 的邻近的解析函数。

早已说过,解析函数的概念只属于精确数学,还可指出,只是在较晚近时,人们才弄清楚泰勒级数在一个节内表示函数的条件。

以前的观点是,只要级数在形式上收敛就够了。不过柯西就已指出,个别点可以是例外。但是,存在着无限制可微函数,它在任何点不能用泰勒级数表示。在这方面,A·普林斯海姆① 1893 年使问题告一段落。他找出除无尽次可微之外,还要加上些什么条件,泰勒级数才的确表示那个函数,即当 n 无限制地增加时,余项才减小到比任何已给量都小。

这里的各种事物有时不是严格地分析清楚的,其所以如此,无疑是与我们曾指出的这样一种观点(在许多书里可以找到)有关:认为在自然界出现的函数都是解析的。

若进一步考察这种观点,就可以看出,它有两个来源:

① "Mathematical papers read at the International Mathematical Congress held in Chicago 1893"("在芝加哥 1893 年国际会议上宣读的论文",纽约,1986 年,第 288—304 页;*Math. Ann.* 第 44 卷,1894 年,第 57—82 页)。

(1) 认为自然界中的函数都是可微的，因而是无尽次可微的。(与此有关的是，认为自然现象都是连续的流行观点。)

(2) 由于不了解普林斯海姆条件的必要性而对它又不加注意。

我当然采取与此相反的观点，认为在自然过程中直接遇到的只是函数带，至于是否宜于把自然过程和精确数学中某些类的函数联系起来，对这个问题的判断就超出我们的感知能力了。

但若说数学家由于数学原因对解析函数有所偏爱，那就是另一回事了。在这里，我愿意附带地略谈数学家对解析函数的偏爱。

这种偏爱的根源是，对于复数 $x = u + iv$，也能界定解析函数。把 $x = u + iv$ 代入泰勒级数，就得到复变函数 $f(x)$ 的展开。倒转来，复函数论也限于研究这样所得到的函数。于是：

一个实变数的解析函数可以拓展到复变数。

具体过程要在函数论里论述。在这里，我只能零星地谈几点。

首先是泰勒级数的收敛域。

函数论指出，对于复数平面的一点 ($x = u + iv$)，泰勒级数的收敛域是一个圆盘，圆盘中心在 x，圆周则经过最近的函数奇点(图 24.9)。级数在圆周上的情形是不确定的，要根据具体情况而定。

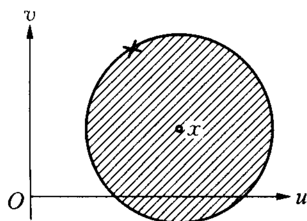


图 24.9

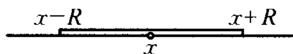


图 24.10

若限于实变数，则这个定理反映出的结论是：收敛域是以所论的点为中心的一个节，但在节的端点，级数的情况不确定(图 24.10)。

我还要说明复变函数论中两个概念：函数元和解析开拓。

我们沿用魏尔斯特拉斯的提法，称函数 $f(x)$ 在一点 a 的泰勒级数

为 $f(x)$的一个函数元。所以它的定义域是以 a 为中心的收敛圆内部。

在许多情况下，我们可以把这个域拓展，方法是在该域内部取另一点 b，然后把原来关于$(x-a)$的幂的级数转化为关于 $x-b$ 的幂的级数。这样所得的新级数的收敛圆可能越过第一个收敛圆延伸到它外面。于是可以在复数平面上新的一个域内界定该函数。这个步骤称为函数的解析开拓，它有时可以连接运用任意多次。在开拓中所得的每一个泰勒级数称为函数的函数元。

归纳起来可以说，原来的级数只直接给出所论解析函数的第一个函数元。但在许多情况下，通过所谓的解析开拓，可以越过原来的区域，并在不断扩大的区域内界定该函数。解析函数的概念包括第一个函数元以及由此出发通过解析开拓所得的新函数元。

更深入地讨论属于解析函数的讲演。对于我们，更需考虑的是事物的哲学方面的问题。首先有这样的问题：关于一个解析函数，我们需要知道些什么，才算了解它的一个函数元（从它可以进一步获得函数的全貌）？

为此，显然必须而且只需知道在 $x=a$ 的泰勒展开式的系数 $f(a)$，$f'(a)$，…这些系数是作为差商的极限值界定的。因此，我们只需了解含有 $x=a$ 的任意一小段。这样，就有如下结论：解析函数的函数元本身，完全决定于曲线 $y=f(x)$ 任意小的一段[①]。

请看，解析函数有着多么特殊的状态，它被整体中任意小的一块所确定。但要注意，这只是在精确数学的意义下，因为只有这时系数 $f(a)$，$f'(a)$，…才能作为差商的极限来界定。

现在，再次考虑自然现象的数学表达问题，我们采取通常的观点，把坐标 x，y，z 看做时间的解析函数

$$x=\varphi(t),\ y=\psi(t),\ z=\chi(t)。$$

---

① ［在一段长时间里，人们以为，只有解析函数才有这样的性质：初始值及其对应的诸导数值就能确定它的整体。E·波雷尔指出，也有非解析函数具备这种性质。请参考 T. Carleman, *Les Fonctions Quasi Analytiques*（卡尔曼：《拟解析函数》，巴黎，1926 年）。］

按照上面所说，就要导致完全的宿命论：任意小一段时间的轨道不可避免地要规定后来一切时刻的轨道。因此必须说：

如果自然界中果真只有解析函数的话，则特殊地，每个小段时间的运动就要用解析函数 $x=\varphi(t)$，$y=\psi(t)$，$z=\chi(t)$ 表示，因而世界的进程必然预先决定于它在任意一小段时间里的状态。

这样的结论是许多人不同意的。因此，就有与此相对立的另一种观点，认为轨道不是用解析函数给出而是用具有解析系数的微分方程给出的。例如

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2}=x,\ y,\ z,\ t,\ \frac{\mathrm{d}x}{\mathrm{d}t},\ \frac{\mathrm{d}y}{\mathrm{d}t},\ \frac{\mathrm{d}z}{\mathrm{d}t}$$

的解析函数。

这种观点是在研究理论力学中形成的。它所导出的不完全是像上面所说的那样严格的结论。例如这样的具解析系数的微分方程的解可能有所谓的“分歧处”，在那里，沿着这条或那条路前进是不确定的。我们可以想想一个一阶微分方程的通解和奇解之间的关系，代表后者的曲线是代表前者的一般积分曲线的包络线(图 24.11)。

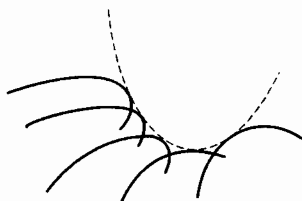


图 24.11

按上述观点，在这样的分歧点，一个点的运动不是完全受一般机械的自然规律制约，可能在那里还有一种非机械的诱因。在巴黎的 J·布森纳斯克(Boussinesq)特别热衷于这个观点。他认为，在解一个具体解析系数的微分方程时，这种分歧点的出现可能是由于在一般机械的世界秩序中介入了一种其他(生物或伦理性质)的力量。他认为，这种分歧点可能用存在着自由意志或者活的物质和死的物质的差异来解释，按照他的想法，在这样的分歧点，活的有机体原则决定了质点的前进道路。参考已发表的 Boussinesq，J.：*Conciliation du véritable déterminisme mécanique avec l'existence de la vieet la*

libertémorale（布森纳斯克：《机械的真实决定性和生命以及精神自由存在性的调和》，巴黎，Gauthier-Villars，1879 年）。

你们了解，对这种观点我将有什么评论。

布森纳斯克谈论的完全是精确数学概念（解析函数确定于一小段等等）。可是，如果我们记得，一切对自然的观察只能达到有限的准确度的话，那么，当我们说，精确数学中的关系仅仅立足于自然时，就是根据一个没有证实的假设。但谈到精确数学时，我们就记得，那里面还含有别的概念，如不连续变数的不连续函数，它们在自然界的数学解释中，同样应能用上①。可是上述整个讨论本身是属于玄学范畴，我们的观察根本不能对它提供任何判断。因此，布森纳斯克的理论在数学上之所以不确切，不是由于那个理论假设了自然现象可以用解析微分方程表达，因而获得错误或不确定的结论，而是由于把上面所说的那个未经证实的假设放在首位，我们可以归纳如下：

所有那类论点之所以脆弱，其原因在于：对精确数学的某些思想和概念有着偏爱，可是对自然的观察却总是只具近似准确性的，它可以多种多样的方式和精确数学相联系②。无论如何，可以争论的问题是：对自然的解释是否本质上要以精确数学为基础，能否通过近似数学的灵巧运用来达到这个目的。

§ 24.7 用有尽三角级数插值法

上面的整个讨论是结合着拉格朗日公式进行的，由此引到泰勒定理，并进一步引到泰勒级数，直到解析函数概念。现在转入用三角级数插值的问题，先考虑有尽三角级数，再考虑无尽的。

假定 $f(x)$是周期函数，周期为 2π，于是 $f(x+2\pi)=f(x)$。

---

① [参考在第 45 页提到的那篇 E·波雷尔报告中关于原子论与数学的说明。]

② 参照像电影那样完全不连续的世界的概念。

我们要用三角级数

$$f(x)=\frac{a_0}{2}+a_1\cos x+\cdots+a_n\cos nx$$
$$+b_1\sin x+\cdots+b_n\sin nx+R$$
$$=\frac{a_0}{2}+\sum_{\nu=1}^{n}(a_\nu\cos\nu x+b_\nu\sin\nu x)+R=\Theta(x)+R,$$

作为这样一个函数的近似表示，其中的余项 R，当它的值无关紧要时，可以略去。有尽级数 $\Theta(x)$ 含有奇数 $(2n+1)$ 个常数。

为了获得级数 $\Theta(x)$ 的表示式，可以采用类似抛物插值的方法，即利用函数的 $2n+1$ 个值 $f(x_0)$，$f(x_1)$，… 来确定那 $2n+1$ 个常数。换句话说，要求 $x=x_1$，x_2… 时，$\Theta(x)$ 的值等于 $f(x_0)$，$f(x_1)$，…，以得到 $\Theta(x)$ 的表示式。

可以立刻仿照拉格朗日公式把所求的 $\Theta(x)$ 写成

$$\Theta(x)=f(x_0)\frac{\sin\frac{x-x_1}{2}\cdot\sin\frac{x-x_2}{2}\cdot\cdots\cdot\sin\frac{x-x_{2n}}{2}}{\sin\frac{x_0-x_1}{2}\cdot\sin\frac{x_0-x_2}{2}\cdot\cdots\cdot\sin\frac{x_0-x_{2n}}{2}}$$
$$+f(x_1)\frac{\sin\frac{x-x_0}{2}\cdot\sin\frac{x-x_2}{2}\cdot\cdots\cdot\sin\frac{x-x_{2n}}{2}}{\sin\frac{x_1-x_0}{2}\cdot\sin\frac{x_1-x_2}{2}\cdot\cdots\cdot\sin\frac{x_1-x_{2n}}{2}}+\cdots$$
$$+f(x_{2n})\frac{\sin\frac{x-x_0}{2}\cdot\sin\frac{x-x_1}{2}\cdot\cdots\cdot\sin\frac{x-x_{2n-1}}{2}}{\sin\frac{x_{2n}-x_0}{2}\cdot\sin\frac{x_{2n}-x_1}{2}\cdot\cdots\cdot\sin\frac{x_{2n}-x_{2n-1}}{2}}。$$

重复应用关于 $\cos\alpha\pm\cos\beta$ 的熟知公式，可以把每个分子化为角的若干倍数的余弦和正弦之和，那就达到了我们原先的目的。于是可以指出：

可以完全仿照拉格朗日多项式那样得到有尽级数 $\Theta(x)$ 的表示

式，而且只要经过形式上的转化，就得

$$\Theta(x)=\frac{a_0}{2}+\sum_{\nu=1}^{n}(a_\nu\cos\nu x+b_\nu\sin\nu x)。$$

在实际应用中，这种插值方式常被采用，不过，在多数情况下，我们假定 x_0，x_1，…，x_{2n} 诸点是等距离的，即把节 x_0，…，$x_0+2\pi$ 分成 $(2n+1)$ 等份，如上面已证明的，这时级数 $\Theta(x)$ 可以用一种直接方法较简捷地导出。等距诸点是

$$x_0,\ x_1=x_0+\frac{2\pi}{2n+1},\ x_2=x_0+2\,\frac{2\pi}{2n+1},\ \cdots,$$

$$x_{2n}=x_0+2n\,\frac{2\pi}{2n+1}。$$

设其对应的函数值是 y_0，y_1，…，y_{2n}，则用以计算 $2n+1$ 个系数的 $2n+1$ 个线性方程

$$y_0=\frac{a_0}{2}+a_1\cos x_0+\cdots$$

$$+a_n\cos nx_0+b_1\sin x_0+\cdots+b_n\sin nx_0,$$

……

$$y_{2n}=\frac{a_0}{2}+a_1\cos x_{2n}+\cdots$$

$$+a_n\cos nx_{2n}+b_1\sin x_{2n}+\cdots+b_n\sin nx_{2n}。$$

为了求得 a_μ 和 b_μ，只需第一次用 $\cos\mu x_\nu$ 乘第 ν 个方程，第二次改用 $\sin\mu x_\nu$ 乘第 ν 个方程 $(\nu=0,1,2,\cdots,2n)$，然后分别相加，再注意 x_ν 和 $x_{\nu+1}$ 之差是 $\frac{2\pi}{2n+1}$，最后可得

$$\sum_{\nu=0}^{2n}\ y_\nu\cos\mu x_\nu=a_\mu\cdot\frac{2n+1}{2},$$

$$\sum_{\nu=0}^{2n}\ y_\nu\sin\mu x_\nu=b_\mu\cdot\frac{2n+1}{2}。$$

其中的计算细节省略了。

这样得到的插值公式，在实际应用中经常使用。同样常用的是，假定有偶数个等距点所得公式①。例如在地磁理论中，就有这样的课题：给出在一条纬线上磁偏角和经度的关系。为此，在 $n+1$ 个等距点观察磁偏角②，按上面公式计算系数 a_μ，b_μ，就得到一个有尽三角级数作为所求的函数。在这里，余项是作为无关紧要来处理的。在具体情况下，这是否恰当，就需由实践者来判断了。

现在特别有意义的是，在上面的公式中，选取无限稠密的 x 值以得到其对应的纵坐标。这时 $\Theta(x)$ 就纯形式地变成无尽三角级数

$$\Theta(x)=\frac{a_0}{2}+\sum_{\mu=1}^{\infty}(a_\mu\cos\mu x+b_\mu\sin\mu x)。$$

可以从我们的公式推得 a_μ，b_μ。

用 Δx_ν 表示两个相邻的点 x_ν 和 $x_{\nu+1}$ 的距离 $\frac{2\pi}{2n+1}$，于是从上面公式得

$$a_\mu=\frac{1}{\pi}\sum_{\nu=0}^{2n}y_\nu\cos\mu x_\nu\cdot\Delta x_\nu,$$

$$b_\mu=\frac{1}{\pi}\sum_{\nu=0}^{2n}y_\nu\sin\mu x_\nu\cdot\Delta x_\nu。$$

因而当 n 无限制地增加，当 Δ_n 趋于零时，就得作为极限值的两个积分

$$a_\mu=\frac{1}{\pi}\int_0^{2\pi}f(x)\cos\mu x\,\mathrm{d}x,$$

$$b_\mu=\frac{1}{\pi}\int_0^{2\pi}f(x)\sin\mu x\,\mathrm{d}x。$$

---

① 在气象学中这些公式以贝塞尔命名。参看他在 *Astronomischen Nachrichten*，1928 年，第 333—348 页的文章："Über die Bestimmung des Gesetzes einer periodischen Erscheinung"（"周期现象规律的确定"）。

② 或者由观察数值，经过插值法以得到这些磁偏角。

以这些系数构成的无尽级数 $\Theta(x)$就是熟知的傅里叶级数。在这里，我们推迟对余项的考察，于是可以把所得结果归结如下：

用无尽傅里叶级数逼近 $f(x)$，可以看做用有尽多个等距纵坐标所得的近似有尽三角级数的极限款[①]。

① 可以和 A·瓦尔特在 *Math. Seminar d. Hamb. Univ.* 上发表的一篇文章相比较，那篇文章通过由殆周期函数转化到纯周期函数以重新处理这个问题。

第二十五章

进一步阐述函数的三角函数表示

§25.1 经验函数表示中的误差估计

上面的讨论还没有涉及余项，即傅里叶级数的收敛性。我们现在来谈这个问题。

我首先指出，关于用到等距纵坐标的有尽三角级数的余项，在所知文献中，还没有一个方便的公式①。但这个级数②却是人们处理自然现象中，遇到作为时间的周期函数时所常用的。在这里，我指出若干经常用到这个工具的学科：气象学（气温及其他气象现象与时间的关系）、声响分析、地磁学（一条纬线上地磁现象和经度的关系）、电技术（交流电强度和时间的关系）等。我要叙述一个来自气象学应用的事实，它和余项估计有关，又密切联系着我们的思路，因而特别值得注意。

以一日 24 小时为横坐标，气温为纵坐标；例如每日观察气温 4 次：午夜 12 时，上午 6 时，中午 12 时，下午 6 时（图 25.1）。然后利用这 4 个数据确定有尽三角级数，以得到气温变化的规律。人们（由于指望余项无关紧要而把它省略掉）得到一个值得注意的结论

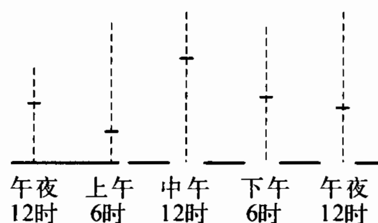


图 25.1

① 参看第 72 页底注①中的建议。

② 以及取偶数个等距离点所得的级数。

是：日出前好几小时，气温达到极低点。这个结论在一切课本中都用了，并且导出最可注意的解释。但在有了自动记录的气温表后，该结论却证实是错误的，真正的最低温度是在紧靠日出之前。实际上气温曲线下降到日出前(例如到上午6时)，然后突然很快上升(图25.2下)。曲线如此下降和突然上升的现象和插值所得结果不一致，因而插值曲线对此显然不能提供正确说明(图25.2上)。应当强调，这里的错误来自略去余项①。于是有如下结论：

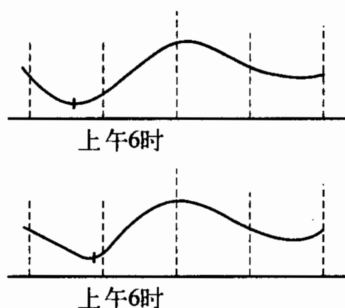


图 25.2

在实际应用中，如何使用这类插值公式，才能得到和观察充分一致的结果，必须在每个具体情况下加以检查处理。上面为气象学证实的例子说明，不加控制地使用插值公式可能导致完全错误的结论。

现在，狄利克莱在 *Crelleschen Journals*(1829年)第四卷的经典论文中已特殊地阐述了傅里叶级数的收敛性及其在表示函数的作用，这理论并已为一切教材和讲演所吸收。人们进行论述时，多半从积分

$$S_n(x)=\frac{1}{2\pi}\int_{-\pi}^{+\pi}f(\xi)\,\frac{\sin\frac{2n+1}{2}(\xi-x)}{\sin\frac{1}{2}(\xi-x)}\mathrm{d}\xi$$

入手，它是由傅里叶级数前 $n+1$ 项的和推导出来的，像我们在第一卷第211页(中译本第224页)所作的那样。

在 $S_n(x)$ 的积分公式中，为了区别于作为固定值的 x，用 ξ 表示

---

① 关于参考文献以及细节，可参考 Ad. Schmidt: *Über die Verwendung trigonometrischer Reihen in der Meteorologie*(施密特：《关于三角级数对气象学的应用》)。Programm des Gymnasium Ernestinum zu Gotha vom Jahre 1894.

积分变数。可以看出，当 n 无限制地增加时，积分极限值是什么。可以证明，在众多情况下，极限值的确就是所要表示的函数 $f(x)$。我省略掉具体的计算，只是指出：

可以把级数的前 $(2n+1)$ 项的和用一个积分表示，并可以证明，若具周期 2π 的函数 $f(x)$ 在节 $0 \leqslant x \leqslant 2\pi$ 里满足所谓狄利克莱条件，则当 $n \to \infty$ 时，该积分必以 $f(x)$ 为极限。狄利克莱条件是：

(1) $f(x)$ 是单值而有界的；

(2) $f(x)$ 是分段连续的，即它只含有有尽多个不连续点；

(3) $f(x)$ 是分段单调的，即只有有尽多个极大极小。

这些充分却不是必要的条件总为我们所说的合理函数所满足，特殊地为（在节 $0 \leqslant x \leqslant 2\pi$ 里）"分段光滑"的函数所满足。这些是具有如下性质的函数 $f(x)$：可以把节 $0 \leqslant x \leqslant 2\pi$ 分为有尽多个子节，在每个子节内部，$f(x)$ 本身和它的导数都连续，而在子节的各端点，当 x 从子节内部向它趋近时，$f(x)$ 和 $f'(x)$ 都趋于有穷值。所以当曲线 $f(x)$ 分段光滑时，它一般地是连续的而且有连续切线，但可能有有尽多个跳跃点和犄角（却没有具纵向切线的尖点）。

此外，$S_n(x)$ 的狄利克莱积分还可以用以考察所给函数 $f(x)$ 在多大程度上可用傅里叶级数中有尽多项近似地表示。为此，只需把积分按适当方式加以估计[①]（参看第一卷第 211 页。中译本第 224 页）。

§ 25.2　通过最小二乘法所得的三角级数插值

在这里，我还回顾傅里叶级数的以下性质：

---

① 同样的方法也适用于采用等距纵坐标所得的有尽级数。

它能够表示简单不连续的函数，即具有简单跳跃点的函数(在节 $0 \leqslant x \leqslant 2\pi$ 里，也只有有尽多个极大点和极小点)。在跳跃点本身，级数的值等于 x 从左和右趋于跳跃点时函数两个极限值的算术中值(图 25.3)，或者，沿用狄利克莱记法

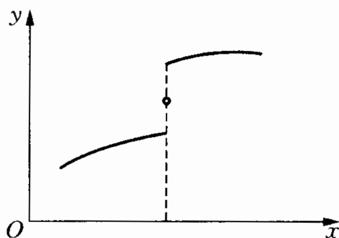


图 25.3

$$\lim_{n\to\infty} S_n(x) = \frac{f(x+0)+f(x-0)}{2}。$$

这只是在本书第一卷中已讨论过的结果，但是，关于傅里叶级数的系数 a_μ，b_μ，也曾给出了另一种推导法，那是在第 63 页提到的论文中贝塞尔用过的。由于下面在另一场合将采用贝塞尔的推导思路，值得回顾一下其主要推导过程。

要想把函数 $f(x)$ 用有尽三角级数

$$\frac{a_0}{2} + a_1 \cos x + \cdots + a_n \cos nx + b_1 \sin x + \cdots + b_n \sin nx$$

表示，我们可以提出这样的课题：确定系数 a_μ 和 b_μ，使它和 $f(x)$ “尽可能好地”近似。宜于采用最小二乘法来解决这个问题。

作函数差

$$f(x) - \frac{a_0}{2} - a_1 \cos x - \cdots - a_n \cos nx$$
$$- b_1 \sin x - \cdots - b_n \sin nx = f(x) - S_n(x),$$

考察用 $S_n(x)$ 表示 $f(x)$ 所产生的误差，并要求误差平方在长度为 2π 的节上的积分

$$\int_0^{2\pi} [f(x) - S_n(x)]^2 \mathrm{d}x$$

有极小值。通过这个容易理解的方式，就可以得到用有尽三角级数

对 $f(x)$的“最好”近似表示。若解出这个极小问题，就可看出，不管 n 是大是小，所得的恰好就是系数 a_μ，b_μ 的傅里叶值，甚至不需要把下标从 1 到 $n(n \geqslant \mu)$ 的项都写出来，也就是可以任意略去其中一些项，所得结果总是一样，于是有以下结论：

用三角级数中任意多项都可以得到函数 $f(x)$的近似表示，其中最好的是通过最小二乘法所得到的。

§25.3　调和分析仪

在已画出曲线 $y=f(x)$ 后，有许多种仪器可以用来机械地算出傅里叶级数里一定数目的系数。这种仪器按英文表达法叫做“调和分析仪”。把一个函数分解为周期项之和的方法，英国人称之为“调和分析”。调和这个词来源于声学，在那里，人们把一个振动分解为尽可能简单的振动，下面介绍的分析仪是科拉迪(Coradi，苏黎世)根据亨里奇(Henrici，伦敦)的思路制作的，当 $\mu=1, 2, \cdots, 6$ 时，它能给出一个 a_μ 和一个 b_μ，至于常数项 a_0，它等于 $\frac{1}{\pi}\int_0^{2\pi} f(x)\mathrm{d}x$，因而是原来曲线下的面积$\frac{1}{\pi}$，可以用平面求积仪求得。

我把这个仪器进一步说明如下：

为了计算傅里叶级数的系数，就要计算积分

$$a_\mu=\int_0^{2\pi}\frac{f(x)\cos\mu x}{\pi}\mathrm{d}x,\ b_\mu=\int_0^{2\pi}\frac{f(x)\sin\mu x}{\pi}\mathrm{d}x$$

(曲线由 $x=0$延伸到$x=2\pi$)。当然，这样的积分一般地可以采用上面提到的机械求积法，甚至可以指出，当只有有尽多次的观察数据时也行[①]，

① 这样，有时候，傅里叶级数就回到上面提到过的用积分界定的系数的有尽三角级数。

在下面这本书里阐述了用图解法求系数的近似值[①]：

Kirsch：*Bewegung der Wärme in den Zylinderwandungen der Dampfmaschine*（基尔施：《蒸汽机圆棒里的热运动》）。莱比锡，1886 年。

另一方面，若曲线已经画出，就可以用一种连续工作的仪器。

为了说明我们的仪器如何工作，需要通过分部积分法把积分化为另一种形式，我们有

$$a_\mu=\frac{1}{\pi}\int_0^{2\pi} y\cos\mu x\,dx=\frac{1}{\pi}\left[\frac{y\sin\mu x}{\mu}\right]_0^{2\pi}-\frac{1}{\pi}\int_0^{2\pi}\frac{\sin\mu x}{\mu}dy,$$

$$b_\mu=\frac{1}{\pi}\int_0^{2\pi} y\sin\mu x\,dx=-\frac{1}{\pi}\left[\frac{y\cos\mu x}{\mu}\right]_0^{2\pi}+\frac{1}{\pi}\int_0^{2\pi}\frac{\cos\mu x}{\mu}dy。$$

若 $y=f(x)$ 不只是分段连续而是连续的，则把上下限代入后，右边两个第一项都等于 0，而分析仪实际上要处理的是

$$\int_0^{2\pi}\sin\mu x\,dy \text{ 和} \int_0^{2\pi}\cos\mu x\,dy。$$

分析仪实现计算这些积分的机制如下：

(1) 它主要含有一个只能平行于 y 轴移动的长方框架（图 25.4）。

(2) 附着在框架上有一个能平行于 x 轴运动的橇，固定在橇上有一根针 St。

---

① ［参考 H. V. Sanden：*Praktische Analysis*（赞登：《实用分析学》，第二版，1924 年，第122—135页），在那里还论述了计算方法。关于计算详情，可参考（第 72 页）已列出的龙格和康尼格的书和波拉克（L. W. Pollack）制的表：*Rechentafeln zur harmonischen Analyse*（《调和分析计算表》，莱比锡，1926 年）。此外，还可以举出弗里泽克（H. Friesecke）、格勒内费尔德（J. Groeneveld）、洛曼（W. Lohmann）、米泽斯（R. von Mises）、波拉泽克-盖林格（H. Pollazek-Geiringer）和齐佩雷尔（L. Zipperer）在 *Zeitschr. f. angewandte Mathematik und Mechanik* 第 2 卷（1922 年），第 3 卷（1923 年）及第 6 卷（1926 年）上的论文。］

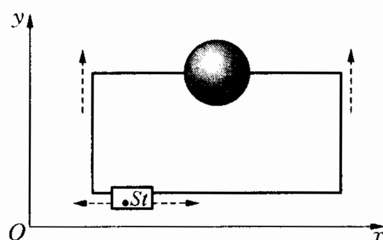


图 25.4

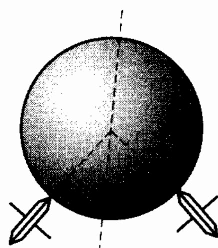


图 25.5

(3) 框架的运动和橇相对于框架的运动合并，橇上那根针就可以沿着任意已给的曲线描绘。

(4) 仪器最重要的部件是一个磨光了的玻璃球。它安装在框架和针相对的那条边上。当框架平行于 y 轴移动 $\mathrm{d}y$ 时，玻璃球在支架上绕平行于 x 的直径转动一个和 $\mathrm{d}y$ 成比例的角(图 25.5)。

(5) 在玻璃球上安装一个带有记录功能的小轮，当玻璃球转动时，它连带地沿一条纬线转动；若纬度是 φ，则转动角和 $\mathrm{d}y\cos\varphi$ 成比例(比例常数和仪器的大小等有关，用 c 表示)。

在玻璃球上有另一个小轮，和前一个位置作 90°角，它记录下 $c\mathrm{d}y\sin\varphi$。两个轮都有记录器，立刻可以读出

$$\int c\cos\varphi\mathrm{d}y \text{ 和} \int c\sin\varphi\mathrm{d}y。$$

(6) 仪器还有一个重要组成部分，它要使这个或那个记录小轮的“纬度”等于 μx(对我们的仪器来说，$\mu=1, 2, \cdots, 6$)，以得到积分

$$\int\cos\mu x\mathrm{d}y \text{ 和} \int\sin\mu x\mathrm{d}y。$$

(7) 为此，在玻璃球上面安上具有适当直径的一个黄铜滑轮，使它的转动带动两个记录小轮。为了使滑轮所转的角和橇上那根针的行程 x 成比例，用银丝绕过它并把它和针相连(图 25.6)。

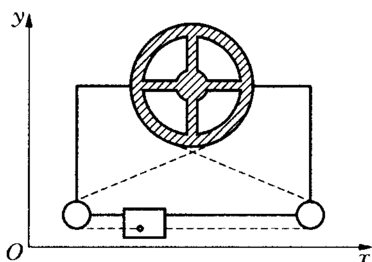


图 25.6

(8) 现在,若选取滑轮直径,使它们和 $1:\frac{1}{2}:\cdots:\frac{1}{6}$ 成比例,则滑轮转角以及记录小轮的"纬度"φ 也和 $1\cdot x$, $2\cdot x$, …, $6\cdot x$ 成比例,这样,从两个小轮就可以读出对应于 $\mu=1, 2, \cdots, 6$ 的积分值,乘以决定于仪器的某个因子。所以,理论是异常简单的。数学上的创造就是通过分部积分把积分变个样。然后根据变了样的积分来制作仪器:首先是作出 dy,然后是分量 $\cos\mu x\,dy$, $\sin\mu x\,dy$,其中主要课题是使记录小轮的"纬度"等于相应的 μx。

§ 25.4 三角级数举例

下面将考察一个函数是如何被它的泰勒级数的部分和所逼近的。为此,我们将充分求助于结合两个例子的曲线图。在这种归纳式的方法中,我宁可局限于初始的论述,因为这样立刻可以引导出一系列在抽象理论中往往难以领会的概念。

我选的第一个例子是这样的 $f(x)$,它的曲线图是由经过 x 轴上 $\pm k\pi(k=0, 1, \cdots)$ 诸点而和 x 轴作 $45°$ 角的线段构成的折线。$f(x)$ 是周期函数,周期是 2π,它在 $0\leqslant x\leqslant 2\pi$ 节上的定义是

$$0\leqslant x\leqslant \frac{\pi}{2} \text{ 时}, f(x)=x;$$

$$\frac{\pi}{2}\leqslant x\leqslant \frac{3\pi}{2} \text{ 时}, f(x)=\pi-x;$$

$$\frac{3\pi}{2}\leqslant x\leqslant 2\pi \text{ 时}, f(x)=-(2\pi-x)。$$

它处处连续，但只是分段解析，因为它的一阶导数在 $(2k-1)\frac{\pi}{2}(k=0, \pm 1, \pm 2, \cdots)$ 有跳跃点。$f(x)$的傅里叶系数 a_μ 全等于 0，因为——我们立刻可以看出——它是奇函数。于是在 $f(x)$的傅里叶级数中，余弦项全部消失。正弦项的系数 b_μ，可以按上面（见第 83 页）给出的公式计算。结果是

$$f(x)=\frac{4}{\pi}\left(\frac{\sin x}{1^2}-\frac{\sin 3x}{3^2}+\frac{\sin 5x}{5^2}-\cdots\right),$$

其中等号成立是因为 $f(x)$满足狄利克莱条件。

在图 25.7 里，画了前两条近似函数的曲线。

$$s_1(x)=\frac{4}{\pi}\cdot\frac{\sin x}{1^2},$$

$$s_2(x)=\frac{4}{\pi}\left(\frac{\sin x}{1^2}-\frac{\sin 3x}{3^2}\right)。$$

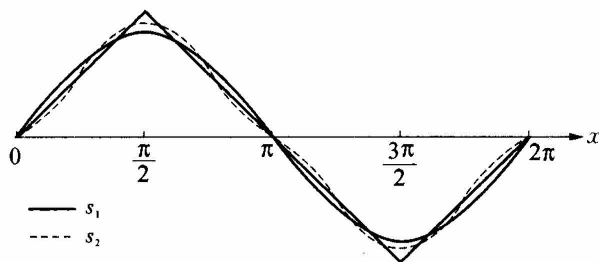


图 25.7

第一条近似曲线是正弦型曲线，在节 $\left(0, \frac{\pi}{2}\right)$ 里和所给曲线相交一次，并且和$(0, \pi)$上的三角形在两个顶点密切。第二条近似曲线在节 $\left(0, \frac{\pi}{2}\right)$ 里和原曲线相交两次，并且像蛇那样在它的两旁来回游动。人们会问，当 n 增大时，近似曲线 $S_n(x)$和原曲线交点数是否无

限制地增加[①]。附带指出，对于我们这个例子，$f(x)$确实如此。

关于这些曲线的纵坐标和 $f(x)$的接近情况，则当 n 增大时，它们的近似程度变得无限制地越来越好，而且对于一切 x 都是如此。至于它们的方向和原曲线的接近，则不能指望在一切点都能实现，例如在 $x=(2k-1)\frac{\pi}{2}(k=0,\pm 1,\cdots)$时，就肯定不能。在那些地方，所有近似函数 $S_n(x)$的导数都等于 0，它们的曲线在那里的切线平行于 x 轴，但原曲线则有犄角。

在考察其他 x 处的方向接近问题之前，我们先考虑第二例：设函数 $g(x)$是具周期 2π 的函数，它在节 $0\leqslant x\leqslant 2\pi$ 里的定义如下：

当 $0<x<\pi$ 时，$g(x)=\frac{\pi}{4}$；

当 $x=0$，或 π，或 2π 时，$g(x)=0$；

当 $\pi<x<2\pi$ 时，$g(x)=-\frac{\pi}{4}$。

在图 25.8 中，画了 $g(x)$的曲线；$g(x)$像前例中的 $f(x)$那样，也是分段解析的，和 $f(x)$不同的是，它只是分段连续的。按狄利克莱定理，$g(x)$可以用它的傅里叶级数表示：

$$g(x)=\frac{\sin x}{1}+\frac{\sin 3x}{3}+\frac{\sin 5x}{5}+\cdots$$

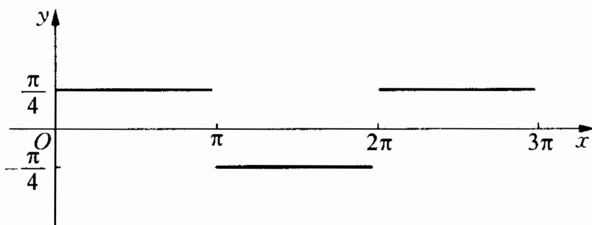


图 25.8

① L·费耶尔在 *Math. Ann.* 第 64 卷(1907 年)第 273—288 页上一篇论文中考察了这个问题。

在图 25.9 中，以不同于图 25.8 的比例尺画出近似函数 $s_1(x)$，$s_6(x)$，$s_{11}(x)$，$s_{16}(x)$ 的曲线。和上例相同，我们首先指出，取的项越多，近似曲线和原曲线的交点数越大①。

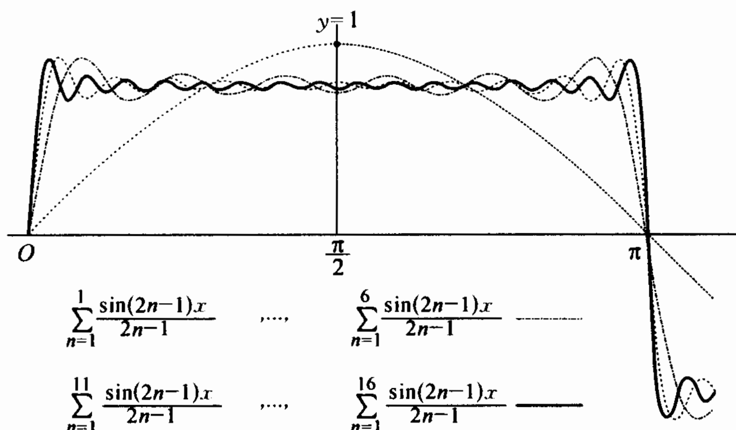


图 25.9

但纵坐标的接近情况又如何呢？

在这里，我们遇到一个细致情节，它引到傅里叶级数在它所代表的函数跳跃点邻近的所谓"非一致收敛"问题。事情是这样的：

例如考虑近似曲线 $s_6(x)$，我们看到，对于离跳跃点不太近的一切 x，它已经很接近 $g(x)$ 的纵标。但在充分靠近跳跃点时，其近似性就越来越差。当 x 例如从右方靠近 0 但又未达到 0 时，原曲线的纵坐标保持在 $\frac{\pi}{4}$，近似曲线却在某处比以前离 $\frac{\pi}{4}$ 更远，然后非常急剧地下降到 0，只要看看 $s_{11}(x)$ 和 $s_{16}(x)$ 就可以察觉，这种情况原则上并非例外。当 n 固定时，只要 x 充分接近非连续点，$S_n(x)$ 就不和 $g(x)$ 接近。

---

① ［在制作图 25.9 时，利用了哥廷根大学数学所的幻灯片。］

若对这情况作出一般性判断（当然不经证明，但可以理解），就是：

对于每个确定值 $x = x_0$，可以取级数足够多项，使近似曲线的纵坐标"满意地"和原曲线的纵坐标一致。x_0 越靠近连续点，就要选取越大的 n，才能达到所希望的满意程度；但若已选定第 n 条近似曲线，则可以令 x_0 充分靠近非连续点，使得该曲线在 x_0 和原曲线达不到任何一致性。

由这个例子所归纳出的事实，可以在理论上叙述为：

满足狄利克莱条件的函数 $f(x)$ 的傅里叶无尽级数，在一切连续点都的确收敛于 $f(x)$，但靠近非连续点时，收敛无限制地变慢①。

人们可能以为，在非连续点，级数根本不收敛。但事实恰恰相反。因为问题中的左右两极限的中值 $\frac{g(0+0)+g(0-0)}{2}$ 等于 0，而第一条近似曲线就已给出这个值。人们之所以有那样的想法，是因为习惯上把函数看成是连续的，因而对相应的级数的收敛性有不符合事实的估计，而实际却恰好不是如此。这里所出现的现象，即在接近一点时，收敛无限制地变慢，称为级数的非一致收敛性。我这里的主要目的是要讲清楚这个概念。

上面已经指出，在靠近一个非连续点时，近似曲线 $s_6(x)$（及其后继者），在非常陡地下降或上升之前，变得相对地远离原曲线。若随着 n 的增加，逐个观察近似曲线的情况，就出现在第一卷，第 214—215 页（中译本第 227—229 页）所讨论的吉布斯现象。它是这样的：设 AB 为跳跃点处连接原曲线上下两端点的线段，则当 n 增大时，曲线 $S_n(x)$ 不是接近 AB 而是超出 A 和 B 同样远的线段 CD。在这里

① ［详情见第一卷，第 211—213 页（中译本第 224—227 页）。］

$$AC = BD = \frac{g(0+0) - g(0-0)}{\pi} \cdot \int_{\pi}^{\infty} \frac{\sin x}{x} \mathrm{d}x^{①}。$$

积分$\int_{\pi}^{\infty} \frac{\sin x}{x} \mathrm{d}x$的值约等于$0.09\pi$，故$AC$和$BD$比$AB$跳跃值约大9%，这结果适用于一切用傅里叶级数表示而有跳跃点的函数$g(x)$。

若再次观察图25.9，但这时集中注意力于不同近似曲线的极大极小的相位差，就会觉得，可以考虑把部分和$S_n(x)$代以它们的算术中值，即

$$S_0(x) = s_0(x), S_1(x) = \frac{s_0(x) + s_1(x)}{2},$$

$$S_2(x) = \frac{s_0(x) + s_1(x) + s_2(x)}{3}, \cdots$$

作为近似函数，费耶尔在*Annalen*上好几篇著名文章中讨论了这些算术中值②。他证明了：对于相当广泛的一类函数，用傅里叶级数中的算术中值$S_n(x)$比部分和$s_n(x)$更能近似地表示函数。在许多情况下，傅里叶级数发散而部分和的中值都收敛于所要代表的函数，特殊地，对于具有图25.9特点的函数，用$S_n(x)$作为近似函数，不出现吉布斯现象。

现在我回到通过近似曲线来接近原曲线方向的问题。对于第一例中的曲线$y=f(x)$，我们曾经指出，除峰点外，它的方向可以用近似曲线逐渐接近。对于第二例中的曲线$g=g(x)$，尽管纵坐标可以接近，其方向则不能。相反地，当n增大时，相对于原曲线，近似曲线摆动得越来越陡。在原曲线一个非连续点邻近，近似曲线的这种表现尤其突出。

从我们的例子，不难看出其原因。

---

① 原书方程右边积分前因子为$\frac{g(x+0) - g(x-0)}{\pi}$，但没有声明$x$是个跳跃点，为明确起见，改成现状。—— 中译者

② [*Math. Ann.* 第58卷(1904年)第51—69页以及第93页所举的文章。]

对于第一曲线，容易画出其导函数曲线。它是由纵坐标等于+1和−1的线段构成，它们交替着在节$\left(0, \frac{\pi}{2}\right)$，$\left(\frac{\pi}{2}, \frac{3\pi}{2}\right)$里。

对于第二曲线，事情较复杂。在这里，一般地导数等于0，只有在跳跃点例外，在那些地方，导数为无穷大。

在第一个例子中，导数曲线本身可以用傅里叶级数

$$f'(x) = \frac{4}{\pi}\left(\cos x - \frac{\cos 3x}{3} + \frac{\cos 5x}{5} - \cdots\right)$$

表示；在跳跃点，它代表在该点两值的中值和原曲线的傅里叶级数

$$f(x) = \frac{4}{\pi}\left(\frac{\sin x}{1^2} - \frac{\sin 3x}{3^2} + \frac{\sin 5x}{5^2} - \cdots\right)$$

比较，可知前者可以从后者通过逐项微分得到。这样，就完全清楚了，近似曲线不但一个个越来越接近原曲线，而且也接近原曲线的方向，只有峰点例外。

第二个例子的情况完全不同。若把原级数

$$g(x) = \frac{\sin x}{1} + \frac{\sin 3x}{3} + \cdots$$

逐项积分，得

$$\cos x + \cos 3x + \cos 5x + \cdots$$

它根本不收敛。由此可见，近似曲线不接近原曲线的方向就不奇怪了。当具有跳跃点的函数用傅里叶级数表示而取函数的导数时，第二个例子中的情况总要出现①。

但若对发散级数

① [关于用傅里叶级数的近似曲线来逼近方向的问题，在下列文献中有准确的阐明：L. Kronecker: *Vorlesungen über Mathematik*（克罗内克：《数学讲义》，第1卷，1894年，第98—99页），特别是 E. W. Hobson: *The theory of funetions of a real variable*（霍布森：《实变函数论》，第二版，剑桥，1926年，第2卷，第639—643页）。后一书对傅里叶级数理论作了深入论述。]

$$\cos x+\cos 3x+\cdots$$

取费耶尔中值，情况就又完全不一样。这些中值，除在跳跃点外，对于一切 x 值都收敛于 $g(x)$ 的导数。于是原来级数的费耶尔算术中值在每个不含跳跃点的节内，既逼近原来的函数又逼近它的导数。这个结论对于一切在节 $0\leqslant x\leqslant 2\pi$ 里，除有尽多个简单跳跃点外，连续而且有连续导数 $g'(x)$（分段光滑）的 $g(x)$，都是对的。其证明见已举出的费耶尔在 *Annalen*（第 58 卷，1904 年）里的论文。

这些就是我想说的关于傅里叶级数的不多的话①。

§ 25.5　切比雪夫关于插值法的工作

全章讨论了关于一元函数的插值及其近似表示。

在结束本章前，我还要特别提到杰出的俄罗斯数学家 P·L·切比雪夫（Tsechebyscheff）的工作。

切比雪夫毕生致力于用简单类型的解析式来近似地表示函数；他成了优秀的近似数学家。

马尔科夫（Markoff）和宗宁（Sonin）把他的著作译成法文，分两卷出版，第一卷含有关于插值理论的众多结果，我在这里特别提出其中 3 篇文章：

(1) “Sur les fractions continues”（“关于连分数”）。Journal de mathématiques pures et appliquées（第二辑）第 3 卷（1858 年）＝论文集 1，第 203—230 页；

(2) “Sur les fonctions qui diffèrent le moins possible de zéro”（“关于与零之差尽可能小的函数”）*Journal de mathématiques pures*

---

① ［关于傅里叶级数的文献，还可以举出《数学百科全书》中伯克哈特（Ⅱ A12）和希尔伯与黎斯（Riesz）（Ⅱ C10）所写的两节；此外，还可以指出下面两书中的有关论述：Courant-Hilbert：*Mathematische Physik* Ⅰ。（柯朗-希尔伯特：《数理物理》，第 1 卷，柏林，1924 年）和 K·克诺普：*Unendliche Reihen*（《无尽级数》，在第 5 页已举出）。］

et appliquées(第 2 辑)第 19 卷(1874 年)＝论文集 2,第 189—215 页;

(3) “Sur le développement des fonctions à une seule variable”(“关于一元函数的发展”)。彼得堡,*Bulletin de I'Acad*,第 1 卷(1859 年)＝论文 1,第 501—508 页。

为了揭示切比雪夫工作的特点,我指出：在我们关于近似表示的论述中,有两个互相交叉的课题。

一方面,对一条曲线给出若干纵坐标,求一个简单的表达式,使它在所给点有这些纵坐标。另一方面,我们采用最小二乘法的思想。

可以通过以下方式把这两方面结合起来：我们要用一个 n 次多项式

$$a+bx+cx^2+\cdots+kx^n$$

代表一条曲线,它含有 $(n+1)$ 个常数。若已给 $(n+1)$ 个观察数据(纵坐标),我们当然就采用拉格朗日插值公式。但已给数据多于 $n+1$ 时,就可以运用最小二乘法思想：我们要求所写下的 n 次抛物线所产生的误差平方之和有极小值。这个课题还可以推广：即不同观察数据可以有不同的权,因而要求(误差)^2×(权)之和达到极小值。对此,切比雪夫利用一种连分法展开,得到一个完整的公式。这样,他的成果的本质也许可以简述如下：

他的第一组工作涉及如下课题：当观察数据多于插值公式所用系数中的未知量个数时,就要求误差平方乘以权之和达到极小值,并据此来确定诸系数。

但切比雪夫并没有单纯地依赖最小二乘法。在另外一些论述中,他不是要求误差平方之和为极小,而是要求所出现的最大误差的绝对值为极小。这项工作出现在他一篇令人惊奇的论文中,其标题是：关于与 0 之差尽可能小的函数。你们即将了解这是什么意思。设已给一个 n 阶多项式,其最高次项是 x^n,问题是,确定其他各项的系数,使得对于在 $+1$ 和 -1 之间的 x,多项式与 0 的差别最小。在

这里，切比雪夫也得到一个简单的最后公式。

他还有第三类的研究工作。

所涉及问题如下：设一条曲线的纵坐标都已给定，但附有一个与横坐标 x 相关的权。所求的是，按照最小二乘法的基本法则，求与所给曲线尽可能好地接近的级数。按照不同观察所得权的分布类型，几乎对实践中用到的级数都获得了结果。其中一个实例是，切比雪夫研讨了曲柄传动装置（所要代表的是曲柄传动装置的规律）问题。我之所以指出切比雪夫关于插值法的工作，一来是因为它很有价值；二来是因为在德国对他的了解还是太少。鲍申格尔在《数学百科全书》第一卷第 6 册（ID3）关于插值法的阐述，是我要向每个从事插值法工作的人推荐的一篇读物，那里也不含有对切比雪夫工作的公正评价。希望有朝一日德国数学界能得到关于他的工作的系统阐述。

第二十六章

二 元 函 数

§26.1 连 续 性

略微回顾迄今为止的讲演就会看出，无论讨论变数 x 或一元函数 $f(x)$，我们始终强调了近似数学和精确数学的对立。和近似数学相联系的则是数学的各种实际应用。

在这里，几何在一定程度上占有核心位置。我们利用几何来引导，使抽象内容较易掌握。事情是这样的：在处理已经画好，并具体地呈现在我们面前的曲线时，首先感觉到的是和近似数学的关系；与此同时，通过几何形象容易揭示抽象观念，间接地使它们能为人所理解。我们用一系列的近似曲线来阐明魏尔斯特拉斯函数时，就是如此[①]。

现在转入二元函数，我们仍然按照这种方式处理。在这里，我们也从阐明精确数学里的概念开始。我的目标是，通过几何图形来讲清楚一些基础关系，而不是像通常那样采取抽象方式。

首先略谈变数 x，y 的区域。要问 x，y 最一般的区域是什么，这个区域什么时候构成连续统等等，就要深入到集合论。我以后再进入这个问题，目前请参考前面已举出的 A·罗森塔尔在《数学百科全书》里的报告。

为简明起见，我们假设函数的定义域是一个圆或者一个长方形，

---

① 也许最终能在一切应用领域按这个意义来利用几何。但在这里，我们不愿深入到这种推测中。

其边平行于坐标轴(图 26.1)。这指的是什么,估计不需要作更多的说明。首先要明确的是,在区域里的每组(x, y)值,对应于一个确定的 z 值。但在每一个具体情况下,必须声明,区域的边界点是否属于它;若边界点属于它,就称区域为闭的,否则称为开的。近来人们把闭区域叫做“区”,开区域叫做“域”①。

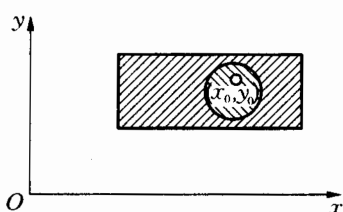


图 26.1

什么时候 $f(x, y)$称为在(x_0, y_0)连续呢? 下面的定义将会得到你们的认可。考虑

$$|f(x, y)-f(x_0, y_0)|,$$

然后说,若 $f(x, y)$满足下面两个条件,它在点(x_0, y_0)就是连续的:第一,它在(x_0, y_0)有唯一的值;第二,已给无论多小的正数 δ,总可以找到一个异于零的 ρ,使得对于一切满足 $(x-x_0)^2+(x-y_0)^2<\rho^2$ 的 x, y,都有

$$|f(x, y)-f(x_0, y_0)|<\delta。$$

[当然我们选取以(x_0, y_0)为中心,$|\rho|$为半径的圆,只是为了简单明确,实质是无论 δ 多小,必有一个含(x_0, y_0)在内的区域,在它内部,总有 $|f(x, y)-f(x_0, y_0)|<\delta$。]

你们看,连续性这个抽象定义本身并无难处。但当我们考虑某些很简单的解析式,例如有理函数

$$z=\frac{2xy}{x^2+y^2}$$

时,却遇到复杂的关系。

---

① “区”和“域”依次是 Bereich 和 Gebiet 的暂译,相当于英文的 region 和 domain。——中译者

这个函数在一切(x, y)连续，只有在$(0, 0)$是例外，因为$x=0$，$y=0$时，分子、分母都等于零，问题是，能否界定函数在$(0, 0)$的值，使它在那里也变成连续。

引进极坐标

$$x = r\cos\varphi,\ y = r\sin\varphi,$$

则当$r \neq 0$时，

$$z = \sin 2\varphi$$

与r无关。

对所给函数的定义作补充，最自然的办法是令$r = 0$时，函数值也是$z = \sin 2\varphi$。

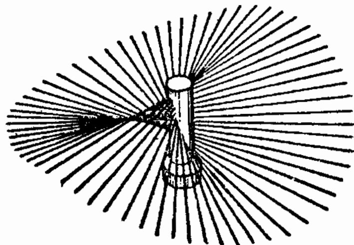


图 26.2

这个函数的几何图像是一个三阶直纹面，在力学中称为"圆柱性面"，如图 26.2（采自 B. St. Ball: *A treatise on the theory of screws*（波尔：《螺杆理论专著》），剑桥，1900 年）所示，它的母线都平行于xy平面并和z轴相交①。

我们的补充假定能使函数z在$x = 0$，$y = 0$处连续吗？显然不能，从图上一眼看出：函数在$(0, 0)$的值并不唯一地确定，它可以是从-1到$+1$的任何值。可是沿着原点的每条路径，它却是连续的。设$y = g(x)$是经原点的一条曲线，其斜率是$\tan\varphi$；则当一点沿曲线趋于$(0, 0)$时，函数z的极限是$\sin 2\varphi$，也就是函数在$x=0$，$y=0$时的值。

---

① 图 26.2 表示曲面$z = \sin 2\varphi$应用于螺杆时的情况；但只有当图中螺杆半径趋于 0 时，得到的才是曲面的真实几何形象。这曲面的"母线"是半线而不是整条直线，因而和通常的直纹面也不相同。"圆柱性面"是 zylindroid（英文 cylindroid）的旧有的中文译名，是从英文字面译来的；实际上曲面本身（如果不考虑应用）与圆柱无关，无论 zylindroid 或圆柱性面都不恰当，易引起误解，姑袭用之。——中译者

假若令函数在(0, 0)有唯一的值,例如0,则连续函数定义中第一条件满足了,但却不能沿每条路径都连续,因而容易看出,连续性定义中第二个条件不能满足。我们说:

多值连续[①]的现象在二元有理函数中已经出现,而在我们连续定义中是排除了的。

接着二元有理函数可能遇到的第一种困难,我还要说明第二种困难。

设函数 $f(x, y)$ 在 (x_0, y_0) 的值唯一确定,再设它沿每条经过 (x_0, y_0) 方向角为 Θ 的直线也连续(图 26.3)。这时 $f(x, y)$,按照我们所述的一般定义是否连续呢? 更详尽些:设对于每个 Θ 和每个无论多小的正数 δ,总有一个正数 ρ_Θ(下标表示与 Θ 有关),使得只要 $r<\rho_\Theta$,就有

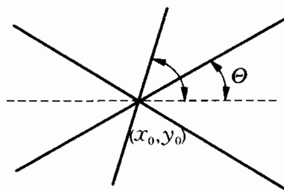


图 26.3

$$|f(x_0+r\cos\Theta, y_0+r\sin\Theta)-f(x_0, y_0)|<\delta。$$

这样,$f(x, y)$ 在 (x_0, y_0) 就一定连续吗?

可以看出,答案不一定是肯定的,即函数沿每个在 (x_0, y_0) 的方向连续,还不一定就意味着它在 (x_0, y_0) 连续。

为了说明这一点,我们先回顾在讨论傅里叶级数的非一致收敛性时的经验(参考前面第37—39页关于一致连续性的讨论)。

满足狄利克莱条件的函数 $f(x)$ 的傅里叶级数在一个非连续点 x_0 的非一致收敛性,或者说收敛变得无限减慢的情况,可以用一条(梯形)曲线来表现。例如,在 x_0 的一个适当选取的邻域里,例如在 x_0 之左,选定一点 x,接着确定级数的项数 n,使其部分和与函数 $f(x)$ 的近似程度至少达到预先规定的要求,然后取 $\frac{1}{n}$ 为纵坐标。这

① stetig-Vieldeutigkeit,指沿不同方向的极限值不全一致。——中译者

样得到的曲线(图 26.4)将能无限制地接近 x_0,但在 x_0 本身,却有异于零的纵坐标,如图 26.4 所示。尽管在 x 接近 x_0 时,收敛无限制地减慢,但却仍然在一切点收敛。

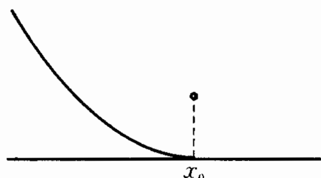


图 26.4

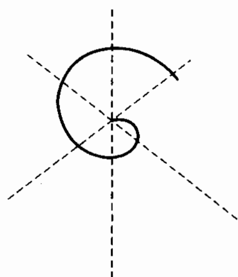


图 26.5

类似情况出现于径向连续。

假定正数 δ 已适当地给定,并且固定不变,对每条经过(x_0, y_0)的直线(方向角为 Θ),计算 ρ_Θ 的最小上界,并在该直线上截取等于这个值的一段以得到从(x_0, y_0)出发的矢径。于是得到一条曲线(图 26.5),它表现不同 Θ 值所对应的连续程度。现在,有可能出现这样一种情况:尽管对应于每个 Θ 值,总有一个异于零的 ρ_Θ,但当靠近某个 Θ 值时,ρ_Θ 却减小到任何固定值 ρ 之下。若曲线靠近某个方向角时,它就任意接近(x_0, y_0),但按径向连续要求,在该直线上应有和(x_0, y_0)距离不等于 0 的点在曲线上,上述情况就要出现。归纳一下:只假定径向连续还不能排除在靠近某个 Θ 值时,ρ_Θ 无限制地接近 0,只是对于该 Θ 值本身,ρ_Θ 才有正值。

在这样的情况下,以(x_0, y_0)为中心显然不能作一个半径为 ρ 的圆,使得在圆内部的一切点,$f(x, y)-f(x_0, y_0)$ 的绝对值总小于 δ。

按下述方法可以构造一个不连续但径向连续的函数。设 Θ 为从原点到点(x, y)的方向角。对于一切异于(0, 0)的(x, y),可以令$\Theta=\operatorname{arc}(x, y)$,考虑按下面条件确定的函数 $F(x)$:

当 $\Theta\neq 0$ 时,$F(x, y)=(n^2-\Theta^2)\sqrt[3]{1:\Theta}$;

当 $\Theta=0$ 时，$F(x, y)=0$；

$F(0, 0)=0$。

这个函数对于每个 (x,y) 有唯一的值。再设 $r=|\sqrt{x^2+y^2}|$，p 是任意正数。则函数

$$z=f(x, y)=r^p F(x, y)$$

在 $(0, 0)$ 只是径向连续。

函数 z 径向连续是可以立刻看出的。因为 $F(x, y)$ 沿一条经过原点的半线是常数，而 r^p 对一切 (x, y) 是连续的。现在，以原点为中心作半径为 δ 的圆，则对于在圆内部的一切点，

$$|f(x, y)-f(0, 0)|=|f(x, y)|<|\delta^p F(x, y)|,$$

而且当 (x, y) 趋于圆周时，$|f(x, y)|$ 可以任意接近 $|\delta^p F(x, y)|$。现在，无论选定 δ 为多么小的正数，我们仍然可以选取 $\Theta=\operatorname{arc}(x, y)$ 如此的小，使得 $|F(x, y)|$，因而 $|\delta^p F(x, y)|$ 要多大就多大。

由此可见，径向连续性和连续性还不是一回事。要从径向连续性达到连续性，必须明确要求，径向一致连续（即对于一切 Θ，能有一个异于零的 δ）。这样就澄清了这个细微的问题。

对于在一个闭区域里连续的函数，有类似一元函数的款中的一些定理，如最大值和最小值都存在，函数要连续地经过两个函数值间的一切值，等等。这些你们自己都能处理。

§26.2 偏导次序的颠倒 $\dfrac{\partial^2 f}{\partial x \partial y} \neq \dfrac{\partial^2 f}{\partial y \partial x}$ 的实例

现在一个问题是：

什么时候连续函数 $f(x, y)$ 是可微的，无限制地可微的，可以按泰勒级数展开的？

首先是关于函数的一阶偏导数

$$\frac{\partial f}{\partial x}=p,\ \frac{\partial f}{\partial y}=q,$$

它们当然不一定存在;我们还是假定它们存在,然后考察更高阶的偏导数。首先是二阶的

$$\frac{\partial^2 f}{\partial x^2}=r,\ \frac{\partial^2 f}{\partial x\partial y},\ \frac{\partial^2 f}{\partial y\partial x},\ \frac{\partial^2 f}{\partial y^2}=t。$$

在这里,我们还需说:

若连续函数 $f(x, y)$在某处不但有一阶偏导数,而且有二阶以及更高阶偏导数,则不但要明确地假定它们的存在,而且特别要考察偏微导是否与次序有关,即例如是否有 $\frac{\partial^2 f}{\partial x\partial y}=\frac{\partial^2 f}{\partial y\partial x}$。

下面我考察上述最后那个问题。我们先假定一切偏导数存在,而且微导次序普遍地可以颠倒,并提出如下问题:

这时 $f(x, y)$是否就可以按泰勒级数

$$f(x, y)=f(x_0, y_0)+\frac{(x-x_0)p_0+(y-y_0)q_0}{1!}$$
$$+\frac{(x-x_0)^2r_0+2(x-x_0)(y-y_0)s_0+(y-y_0)^2t_0}{2!}$$
$$+\cdots$$

$\left(s=\frac{\partial^2 f}{\partial x\partial y}=\frac{\partial^2 f}{\partial y\partial x}\right)$展开?

根据以前的考察,我们必须如此回答:

偏导数存在以及形式上写下的级数收敛,并不保证泰勒级数展开式就适用,要这个展开式能用,普林斯海姆条件必须满足,即当 n 增大时,余项必须任意地接近 0。只有当这些都满足时,我们才说 $f(x, y)$首先在收敛域内,是含变数(x, y)的解析函数。

我们还要补充一点:由此可见,还需要作大量的假设,$f(x, y)$

才能是解析函数。

现在让我们稍微看看应用中的情况。我们试问：在力学和物理学中用到二元或多元函数时，情况是如何的呢？

在应用中，通常的说法是，人们把每个函数都看做解析的，因而相信可以把级数从一阶、二阶或三阶断开，“因为这给出充分近似的表示”。

我们的观点与此相反。

当然，上述那样的表述方式和我们所阐明的解析函数的严格概念是不相容的。在那里，人们实际上从头就是在近似数学领域里活动，并且贸然假定可以用一次、二次或三次多项式来充分近似地表示问题中的函数。因此，那种通常提法的主要依据是：假定在应用中出现的函数可以充分近似地用一次、二次、三次多项式表示。至于它们是否“真的”把函数作为解析函数来处理，我们无法作出判断，因而我不能对此作出评论。

问题在于：在应用中，是否真的仅仅涉及上述近似表示的存在性；特殊地，在实际中（不管导数在那里的意义如何确定）是否总有 $\frac{\partial^2 f}{\partial x \partial y}=\frac{\partial^2 f}{\partial y \partial x}$？或者整个假设只是来自流行的习惯而没有深入到事物本质呢？

我必须说：在实践中有充分的事例表明，有时不能假定

$$\frac{\partial^2 f}{\partial x \partial y}=\frac{\partial^2 f}{\partial y \partial x}。$$

为了把问题说清楚，我们先问：

在精确数学里，人们是在什么条件下，证明 $\frac{\partial^2 f}{\partial x \partial y}=\frac{\partial^2 f}{\partial y \partial x}$ 的？下面我们将要通过对曲面图形的观察，举出这些条件不满足的实例。

$\frac{\partial^2 f}{\partial y \partial x}$ 的定义是这样的：

首先有

$$\frac{\partial f}{\partial y}=\lim_{k\to 0}\frac{f(x,\ y+k)-f(x,\ y)}{k}=q(x,\ y),$$

和
$$\frac{\partial q}{\partial x}=\frac{\partial^2 f}{\partial y\partial x}=\lim_{h\to 0}\frac{q(x+h,\ y)-q(x,\ y)}{h}。$$

合并两等式,则

$$s'=\frac{\partial^2 f}{\partial y\partial x}$$
$$=\lim_{h\to 0}\left(\lim_{k\to 0}\frac{f(x+h,\ y+k)-f(x+h,\ y)-f(x,\ y+k)+f(x,\ y)}{hk}\right),$$

因此

$\frac{\partial^2 f}{\partial y\partial x}$ 是对右边的商先令 k,后令 h 趋于 0 的结果。

同样,我们有

$$s=\frac{\partial^2 f}{\partial x\partial y}$$
$$=\lim_{k\to 0}\left(\lim_{h\to 0}\frac{f(x+h,\ y+k)-f(x+h,\ y)-f(x,\ y+k)+f(x,\ y)}{hk}\right)。$$

因此,若把取极限的次序颠倒,即先令 h,后令 k 趋于 0,就得到另一个偏导数。

两个二阶偏导数是否相等决定于两个极限过程可否颠倒。我们现在利用中值定理证实,在通常情况下,的确 $s=s'$。

我们有

$$p(x,\ y)=\lim_{h\to 0}\frac{f(x+h,\ y)-f(x,\ y)}{h}。$$

假定 $p(x,\ y)$在$(x,\ y)$和在$(x,\ y)$的一个邻域 U 里(有穷而且唯一)存在,则利用中值定理,在适当选取 h 的情况下,

$$f(x+h,\ y)-f(x,\ y)=hp(x+\theta h,\ y),$$

其中 $0<\theta<1$。

现在假定 p 在 U 里对 y 可微。这时可以再次利用中值定理：在适当选取 k 的情况下，

$$p(x+\theta h, y+k)-p(x+\theta h, y)=ks(x+\theta h, y+\eta k),$$

其中 $0<\eta<1$。或者，合起来

$$f(x+h, y+k)-f(x, y+k)-f(x+h, y)+f(x, y)$$
$$=hks(x+\theta h, y+\eta k)^{①}。$$

以 k 除等式两边，就得

$$\frac{f(x+h, y+k)-f(x+h, y)}{k}-\frac{f(x, y+k)-f(x, y)}{k}$$
$$=hs(x+\theta h, y+\eta k)。$$

现在假定 q 在 (x, y) 和 U 中存在，则令 $k\to 0$ 时，左边趋于

$$q(x+h, y)-q(x, y)。$$

在右边，若再假设在邻域 U 里，对于常值 x，s 是 y 的连续函数，则

$$\lim_{k\to 0}s(x+\theta h, y+\eta k)=s(x+\theta h, y)。$$

因此

$$\frac{q(x+h, y)-q(x, y)}{h}=s(x+\theta h, y)。$$

最后，还假设在 (x, y)，s 对于 x 连续。这时，令 h 趋于 0，即得

$$\lim_{h\to 0}\frac{q(x+h, y)-q(x, y)}{h}=s(x, y)。$$

但左边极限无非就是 $s'(x, y)$，于是 $s'(x, y)$ 存在而且等于 $s(x, y)$。

在我们的证明里，不但必须假定 p，q，s 在 (x, y) 以及 (x, y) 的一个邻域里存在而且还要求 s 满足一定的连续条件。若略去上面这

---

① 偏导换序的这个充分条件，可参阅菲赫金哥尔茨（Г. М. Фихтенгольц）著《微积分学教程》第 1 卷，第二分册第 388—389 页。—— 中译者

个或那个条件，就不再能指望s'——即使存在——等于s①。

我们即将看到，在日常遇到的曲面②中，就有 $s \neq s'$ 的例子。

设有两个半径相同的圆柱形相交所形成的“交叉拱顶”(图 26.6)。选取两柱面交线的最高点作为坐标原点，z 轴向上，然后计算 s 和 s'在原点的值。为了得到不被拱顶对称性掩盖的有用结果，设想整个拱顶相对于坐标轴转动 α 角③，$0<\alpha<45°$，那样，在 xy 平面上的投影就像图 26.7 所示的那样。

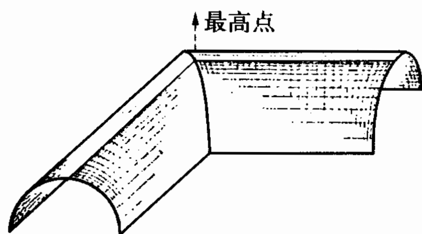


图 26.6

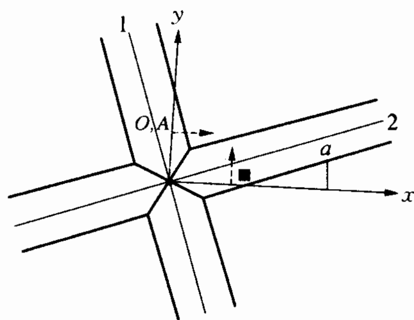


图 26.7

① [关于上面所论的颠倒偏导次序的问题，H · A · 许瓦兹第一次作了奠基性处理。见他的论文：“Über ein vollständiges System von einander unabhängiger Voraussetzungen zum Beweise des Satzes: $\frac{\partial}{\partial y}\left(\frac{\partial f(x, y)}{\partial x}\right)=\frac{\partial}{\partial x}\left(\frac{\partial f(x, y)}{\partial y}\right)$”(“关于证明定理$\frac{\partial}{\partial y}\left(\frac{\partial f(x, y)}{\partial x}\right)=\frac{\partial}{\partial x}\left(\frac{\partial f(x, y)}{\partial y}\right)$中的一组完备而且互相独立的假设”) (*Ges. Math. Abh.* (《数学论文集》，第 2 卷，1890 年，第 275—284 页)。——内德(L. Neder)研究了三元和多元函数以及二阶以上的偏导数(*Math. Zeitschr*，第 24 卷(1926 年)，第 759—772 页)]。

② 或者(因为我们看到的不是精确的曲面)代表日常所遇见曲面的理想曲面。

③ 在这里，无形中假定了两个圆拱的母线互相正交，而且原来是沿 x 轴和 y 轴方向。——中译者

现在计算在拱顶上每个点的 p 和 q，作为差商的极限值。在原点，当然 $p=0, q=0$，因为两个圆拱沿最高线的切面是水平面；故 $p(0,0)=q(0,0)=0$。

那么 $s(0,0)$ 和 $s'(0,0)$ 又如何呢？我们把它们看做下面的极限

$$s=\lim_{k\to 0}\frac{p(0,k)-p(0,0)}{k},\ s'=\lim_{h\to 0}\frac{q(h,0)-q(0,0)}{h}。$$

为方便起见，设 h 和 k 的值都是正的。这样，点 $(0,k)$ 在圆拱 1 上，而 p 就是圆拱 1 沿 x 方向上升或下降的斜率。于是 $p(0,k)$ 是负的。另一方面，点 $(h,0)$ 在圆拱 2 上；q 是圆拱 2 沿 y 方向上升或下降的斜率。故 $q(h,0)$ 是正的。由这个定性的观察可知：

s 是作为负值差商的极限，s' 是作为正值差商的极限得到的。因此若 s 和 s' 不是都等于 0，它们就不能相等。

我们现在利用公式来计算这两个极限。

当两个圆拱处于原来的位置时，它们的母线平行于 x，y 两坐标轴，其方程是①

圆拱 2：$z=C_0-C_1y^2-\cdots$

圆拱 1：$z=C_0-C_1x^2-\cdots$

因为 $p(0,0)=0$，$q(0,0)=0$，其中一次项不出现；为简洁起见，高于二次的项不再写出，因为它们不影响推导结果。

现在，令圆拱位置如图 26.7 所示，并用Ⅰ，Ⅱ，Ⅲ，Ⅳ表示整个交叉拱顶的不同部位(图 26.8)，就得：对于Ⅰ和Ⅲ(圆拱 1)：

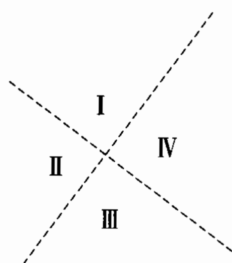


图 26.8

$$z=C_0-C_1(x\cos\alpha+y\sin\alpha)^2-\cdots$$

对于Ⅱ和Ⅳ(圆拱 2)：

① 原文方程中最后是正号，现改的和下面方程一致。——中译者

$$z = C_0 - C_1(-x\sin\alpha + y\cos\alpha)^2 - \cdots$$

于是，

对于Ⅰ和Ⅲ

$$p = -2C_1 x\cos^2\alpha - 2C_1 y\sin\alpha\cos\alpha - \cdots$$
$$q = -2C_1 x\sin\alpha\cos\alpha - 2C_1 y\sin^2\alpha - \cdots$$

对于Ⅱ和Ⅳ

$$p = -2C_1 x\sin^2\alpha + 2C_1 y\sin\alpha\cos\alpha - \cdots$$
$$q = +2C_1 x\sin\alpha\cos\alpha - 2C_1 y\cos^2\alpha - \cdots$$

根据这些公式，对于交叉拱顶，在原点的确有 $p=0$，$q=0$。可是沿两个圆拱的交线，p 和 q 都不连续；因两圆拱相遇处有折痕。

现在对原点计算 s。由Ⅰ中的 $p(0, k)$，若略去展开式的高次项，得

$$s = \lim_{k\to 0}\frac{-2C_1 k\sin\alpha\cos\alpha}{k} = -C_1\sin 2\alpha。$$

另一方面，由Ⅳ，略去高次项后，得

$$s' = \lim_{h\to 0}\frac{+2C_1 h\sin\alpha\cos\alpha}{h} = +C_1\sin 2\alpha。$$

因此，一般地 $s \neq s'$，而且只有当坐标轴和两拱面母线平行时，$s = s'$。

若在Ⅱ区和Ⅳ区计算(0,0)邻近的 s 值而在Ⅰ和Ⅲ区计算(0,0)邻近的 s'值，则略去高次项后，得

$$s = +C_1\sin 2\alpha,\ s' = -C_1\sin 2\alpha。$$

由此可知 $s(x, y)$和 $s'(x, y)$在(0, 0)都不连续。

你们看，这个例子既不抽象，也不难，它们是很平常的。例如它可以推广到雨伞的曲面，只要伞是直竖的，其左右对于 xz 平面和 yz① 平

① 原书作 xy。——中译者

面不对称，就会在其顶点出现 $s \neq s'$。我们也许可以写出以下结论：

上述例中的那种曲面上所出现的 s 和 s' 不相等的情况是很平常的，我们之所以没有想到，是因为我们很少留心具体事物所提供的情况（而总是不假思索地接受教科书中的结论）。

§26.3 用球函数级数近似表示球面上的函数

上面从精确数学角度谈论了二元函数。我现在转到从近似的角度来谈，就像从精确数学角度讨论了函数 $f(x)$ 之后，接着就讨论插值法和近似表示那样。

人们将怎样近似地表示函数 $f(x, y)$ 呢？

最易考虑到的是，用次数递增的，含 x，y 的齐次多项式来逼近 $f(x, y)$

$$z = f_0 + f_1 + f_2 + \cdots$$

可以指出，这是在力学和物理学中常用的方法。

另一种方法是，试用与三角级数相应的拉普拉斯球函数来逼近函数。

在这里，我略过第一种方法，只对拉普拉斯球函数略加阐述，主要是说明它们的本质及其在表示球面上的函数中的作用①。

这里的问题联系着三角级数的一种推广。为了说明这一点，我们必须把以前三角级数中的辐角 x 换成 ω，于是得 $f(\omega)$ 的三角级数表示

$$f(\omega) = \frac{a_0}{2} + a_1 \cos\omega + \cdots$$

① 对下面的思路概述如下：（一元）周期函数实际上是幺圆（半径等于 1 的圆）周上的函数，向高一维推广，是幺球面上的函数。作者对周期函数的三角级数近似表示在形式上加以改变，然后通过拉普拉斯方程这个纽带，从这种新表示形式自然地过渡到用球函数近似地表示幺球面上的函数。——中译者

$$+a_n\cos n\omega+\cdots+b_1\sin\omega+\cdots+b_n\sin n\omega+\cdots$$

下面我们将对此在形式上作些改变。

(1) 我们设想 ω 是圆心角，z 是一个在半径为 1 的圆周上、周期为 2π 的函数(图 26.9)。

(2) 这个函数 $f(\omega)$ 将要通过像 $a_n\cos n\omega+b_n\sin n\omega$ 那样一些成双的项近似地表示，其间的规律，我们引进直角坐标后用另一种形式来表达。

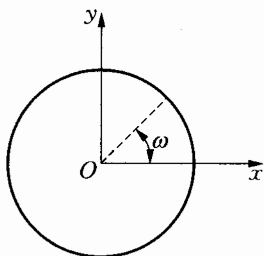


图 26.9

由 $\cos\omega=x$，$\sin\omega=y$，得

$$\cos\omega+\mathrm{i}\sin\omega=x+\mathrm{i}y,$$

又根据 A·棣莫佛公式，

$$\cos n\omega+\mathrm{i}\sin n\omega=(x+\mathrm{i}y)^n,$$

$$\cos n\omega-\mathrm{i}\sin n\omega=(x-\mathrm{i}y)^n。$$

由此(通过很明显的途径)可知：

(3) $\cos n\omega$ 和 $\sin n\omega$ 可以用含 x, y 的 n 次多项式表示，$a_n\cos n\omega+b_n\sin n\omega$ 也是如此。所有这样的多项式 F 都满足简单的微分方程

$$\Delta F=\frac{\partial^2 F}{\partial x^2}+\frac{\partial^2 F}{\partial y^2}=0,$$

因为 $(x+\mathrm{i}y)^n$ 和 $(x-\mathrm{i}y)^n$ 就都满足它。最重要的是，这个微分方程代表着这些多项式的特征。

这句话是指：

(4) 一个满足方程 $\Delta F=0$ 的齐 n 次多项式总是和某个 $a\cos n\omega+b\sin n\omega$ 联系着。

这可以通过简单计数推知：一个齐 n 次多项式有 $n+1$ 个系数。其相应的 ΔF 是一个齐 $(n-2)$ 次多项式，有 $(n-1)$ 个系数，而 ΔF 应等于 0。因此，那 $(n+1)$ 个系数要满足 $(n-1)$ 个条件：有两个系数还

可以任意选取(对应于 $a\cos n\omega + b\sin n\omega$ 中的 a 和 b)。

若用 F_0，F_1，…，F_n 依次代表所说的从 0 次到 n 次的多项式，我们就有结论：

函数 $f(\omega)$的通常的三角级数表示可以写成升幂的含 x，y 的齐次多项式 F_ν 的级数展开式，其中各多项式 F_ν 都满足特殊的方程 $\frac{\partial^2 F_\nu}{\partial x^2} + \frac{\partial^2 F_\nu}{\partial y^2} = 0$。得到展开式后，再令

$$x = \cos\omega,\ y = \sin\omega。$$

在这里，完全没有考虑所处理的级数是否有尽，而仅仅是纯粹形式上运用了各项的转化规律。当然还存在着计算那些待定常数 a_n，b_n 的技术问题。

在我对普通的三角函数如此描述之后，它就容易被推广到多元函数。我们只考虑高一维的情况。

取半径为 1 的球面，设 x，y，z 是它上面点的直角坐标。引进球面坐标 φ,θ,其中 φ 表示经度,θ 表示极纬度[①] $0 \leqslant \theta \leqslant \pi$，就有公式(图 26.10)

$$x = \sin\theta\cos\varphi,$$
$$y = \sin\theta\sin\varphi,$$
$$z = \cos\theta。$$

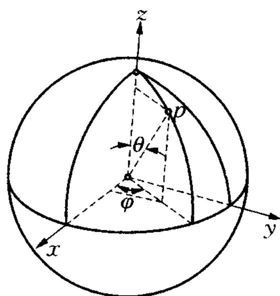


图 26.10

现在设已给球面上的一个函数$f(\theta,\ \varphi)$，我们要看看，是否类似地可以用含 x，y，z 的齐 ν 次($\nu=0$，1，2，…)多项式 F_ν 的级数表示，而现在这些多项式则满足方程

$$\Delta F_\nu = \frac{\partial^2 F_\nu}{\partial x^2} + \frac{\partial^2 F_\nu}{\partial y^2} + \frac{\partial^2 F_\nu}{\partial z^2} = 0。$$

① 极纬度是由北极(0，0，1)量起的纬度。——中译者

这样

$$f(\theta, \varphi) = F_0 + F_1 + \cdots + F_n + \cdots$$

这个级数，如果无尽，是否收敛，那些有尽级数是否能适当近似地表示函数 $f(\theta, \varphi)$，本身都是问题，这些都暂且不谈。

一个(拉普拉斯)球函数首先是满足微分方程 $\Delta F=0$ 的一个含 x, y, z 的齐 n 次多项式。

稍迟些，一个齐 n 次球函数将只用来代表把 x, y, z 作为 θ, φ 的函数代入后所得的表达式。

考虑一下，球函数 F_n 中还有几个待定常数。一个任意的 $F_n(x, y, z)$ 有 $\frac{(n+1)(n+2)}{2}$ 个常数(利用归纳法，由 F_1 有 3 个，F_2 有 6 个可以推得)。ΔF_n 是齐 $(n-2)$ 次多项式，有 $\frac{(n-1)n}{2}$ 个常数。令 ΔF_n 恒等于零，就得到 $\frac{(n-1)^n}{2}$ 个条件；因此还剩下

$$\frac{(n+1)(n+2)-(n-1)n}{2} = \frac{4n+2}{2} = 2n+1$$

个常数是任意的，于是有定理：

最一般的 n 次球函数含有 $2n+1$ 个未定参数；这些参数满足一个线性方程组，因为由 $\Delta F_n=0$ 得到的只是关于 F_n 的系数的方程。

试考察 $n=2$ 的款。最一般的齐二次多项式是

$$F_2 = a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{23}yz + 2a_{31}zx + 2a_{12}xy。$$

故 $\Delta F_2=0$ 就是

$$2(a_{11} + a_{22} + a_{33}) = 0。$$

所以 6 个系数要满足一个线性条件，因而还有 5 个系数是任意的，和定理相符。

我们给出从 0 次到四次的这些多项式(见第 119 页表)。一般作

法是这样的：对于 $n=0, 1, 2, 3, 4, \cdots$ 各款，分别给出用球面坐标表示的 $2n+1$ 个特殊的球函数（参看本页从 0 次到四次的球函数表）；一般 n 次球函数就是这 $(2n+1)$ 个特殊球函数的线性组合，其系数是任意的。

F_0	1				
F_1	$\cos\theta$	$\left.\begin{matrix}\sin\theta\cos\varphi\\ \sin\theta\sin\varphi\end{matrix}\right\}1$			
F_2	$3\cos^2\theta-1$	$\left.\begin{matrix}\sin\theta\cos\varphi\\ \sin\theta\sin\varphi\end{matrix}\right\}$ $\cos\theta$	$\left.\begin{matrix}\sin^2\theta\cos 2\varphi\\ \sin^2\theta\sin 2\varphi\end{matrix}\right\}1$		
F_3	$5\cos^3\theta-$ $3\cos\theta$	$\left.\begin{matrix}\sin\theta\cos\varphi\\ \sin\theta\sin\varphi\end{matrix}\right\}$ $(5\cos^2\theta-1)$	$\left.\begin{matrix}\sin^2\theta\cos 2\varphi\\ \sin^2\theta\sin 2\varphi\end{matrix}\right\}\cos\theta$	$\left.\begin{matrix}\sin^3\theta\cos 3\varphi\\ \sin^3\theta\sin 3\varphi\end{matrix}\right\}1$	
F_4	$35\cos^4\theta-$ $30\cos^2\theta+3$	$\left.\begin{matrix}\sin\theta\cos\varphi\\ \sin\theta\sin\varphi\end{matrix}\right\}$ $(7\cos^3\theta-3\cos\theta)$	$\left.\begin{matrix}\sin^2\theta\cos 2\varphi\\ \sin^2\theta\sin 2\varphi\end{matrix}\right\}$ $(7\cos^2\theta-1)$	$\left.\begin{matrix}\sin^3\theta\cos 3\varphi\\ \sin^3\theta\sin 3\varphi\end{matrix}\right\}\cos\theta$	$\left.\begin{matrix}\sin^4\theta\cos 4\varphi\\ \sin^4\theta\sin 4\varphi\end{matrix}\right\}1$

我建议读者把上表中第一列的球函数计算一下①。其余的只需把第一列的诸项对 $\cos\theta$ 微导，去掉系数中某些数值因子，然后乘上表中所列出的相应因子即可得到②。

最一般的球函数 F_4 就是表中末行 9 个特殊球函数 F_4 的线性组合，其系数可以任意。

现在考虑一个课题：已给函数 $f(\theta, \varphi)$，按照最小二乘法的思路，用 0 次到 4 次的球函数来近似地表示它。这就是：令

$$f(\theta, \varphi)=F_0+F_1+F_2+F_3+F_4+\text{余项},$$

① 对已给出的多项式验算起来当然容易。例如回到直角坐标，就可以把 $F_3=5\cos^3\theta-3\cos\theta$ 写成 $F_3=5z^3-3z(x^2+y^2+z^2)$，然后验证 $\Delta F_3=0$。

② 对 $\cos\theta$ 取表中第一列各行的 ν 阶导数（去掉某些数值因子），就得到同行第 $\nu+1$ 列花括号后的因子，$\nu=1, 2, 3, 4$。——中译者

并要求误差平方在幺球面上的积分

$$\int(f-F_0-F_1-\cdots-F_4)^2\cdot \mathrm{d}o$$

为极小,其中 do 表示幺球面的面积微元。在一些书里,我们看到的多半只是把一个函数按球函数无尽级数展开,只有在实践中才遇到用有尽级数作近似表示的问题。我们的课题是:计算 $F_0+F_1+\cdots+F_4$ 中的 25 个系数,使极小值的要求得到满足。

现在看来问题是要从具一般结构的 25 个线性方程求 25 个未知数。我们即将看到,事情并不那么坏,因为由每个方程可以解出一个未知数。

为了计算便于进行,我们采用总和记号并且按通常办法用 $P_n(\cos\theta)$表示表中第一列的球函数,表中其余球函数可以写作

$$\left.\begin{matrix}\sin^{\nu}\theta\cos\nu\varphi\\ \sin^{\nu}\theta\sin\nu\varphi\end{matrix}\right\}P_n^{(\nu)}(\cos\theta),$$

其中 $P_n^{(\nu)}(\cos\theta)$表示 P_n 对 $\cos\theta$ 的 ν 阶导数(去掉某个数值因子)。于是整个级数是

$$\begin{aligned}&f(\theta,\ \varphi)\\ &=\sum_{n=0}^{4}\sum_{\nu=0}^{4}[(a_{n,\nu}\sin^{\nu}\theta\cos\nu\varphi+b_{n,\nu}\sin^{\nu}\theta\sin\nu\varphi)P_n^{(\nu)}(\cos\theta)],\end{aligned}$$

我们把它简写成

$$\sum\sum(a_{n,\nu}\Phi_{n,\nu}+b_{n,\nu}\Psi_{n,\nu})。$$

当我们开始考察误差平方积分

$$\int\{f(\theta,\ \varphi)-\sum\sum(a_{n,\nu}\Phi_{n,\nu}+b_{n,\nu}\Psi_{n,\nu})\}^2\mathrm{d}o$$

时,则由于球函数的正交性,含诸 a, b 的二次项大为简化。

所谓正交性是这样的:用 F', F''表示 25 个特殊球函数中不同

的两个，则$\int F'F''\mathrm{d}o=0$，或者，特殊地

$$\int \Phi'\Phi''\mathrm{d}o=0,\ \int \Phi'\Psi''\mathrm{d}o=\int \Phi''\Psi'\mathrm{d}o=0,$$

$$\int \Psi'\Psi''\mathrm{d}o=0\ ①。$$

因此，函数Ω(误差平方积分，令它的值为极小将给出所要求的系数a，b)变得比我们预料的要简单得多，它化为以下形式：

$$\Omega=\int f^2\mathrm{d}o-2\sum\int f\cdot F\cdot \mathrm{d}o+\sum\int F^2\mathrm{d}o,$$

其中总和的范围是所有上面给出的特殊球函数。用Φ和Ψ表示，Ω可以写成

$$\Omega=\int f^2\mathrm{d}o-2\sum\sum\left[a_{n,\nu}\int f\Phi_{n,\nu}\mathrm{d}o+b_{n,\nu}\int f\Psi_{n,\nu}\mathrm{d}o\right]$$
$$+\sum\sum\left[a_{n,\nu}^2\int \Phi_{n,\nu}^2\mathrm{d}o+b_{n,\nu}^2\int \Psi_{n,\nu}^2\mathrm{d}o\right]。$$

为了使它有极小值，令它对于$a_{n,\nu}$和$b_{n,\nu}$的偏导数等于0。于是得

$$\frac{1}{2}\frac{\partial\Omega}{\partial a_{n,\nu}}=-\int f\Phi_{n,\nu}\mathrm{d}o+a_{n,\nu}\int \Phi_{n,\nu}^2\mathrm{d}o=0,$$
$$\frac{1}{2}\frac{\partial\Omega}{\partial b_{n,\nu}}=-\int f\Psi_{n,\nu}\mathrm{d}o+b_{n,\nu}\int \Psi_{n,\nu}^2\mathrm{d}o=0。$$

其中积分范围都是整个球面。我们可以看到，所得的25个方程每个只含一个未知数。我们立刻得到

$$a_{n,\nu}=\frac{\int f\Phi_{n,\nu}\mathrm{d}o}{\int \Phi_{n,\nu}^2\mathrm{d}o},\ b_{n,\nu}=\frac{\int f\Psi_{n,\nu}\mathrm{d}o}{\int \Psi_{n,\nu}^2\mathrm{d}o}。$$

---

① [可以和三角级数中类似积分相比较：即$\mu\neq\nu$时的积分$\int_0^{2\pi}\cos\mu x\ \cos\nu x\mathrm{d}x$，$\int_0^{2\pi}\sin\mu x\sin\nu x\mathrm{d}x$以及$\mu\neq\nu$或$\mu=\gamma$时的积分$\int_0^{2\pi}\sin\mu x\cos\nu x\mathrm{d}x$。]

于是我们的课题已经解决。我们补充一项一般性结果：

这样解答方法的妙处在于(和傅里叶级数完全类似)；令误差平方的积分极小所得到的每个 $a_{n,\nu}$ 和 $b_{n,\nu}$ 的值同用以对 f 作近似表示的别的下标 n，ν 毫不相干。所以无论级数展开如何延伸，一经计算出的 $a_{n,\nu}$ 就保持不变，因此，已算出用 0 次到四次球函数近似表示中的 25 个系数，当加上五次球函数后，它们仍然适用，只需添上 11 个新系数即可。

§26.4 球函数在球面上的值分布

以上都只是纯形式的处理。为了对这些只是形式上界定的球函数的本质有粗略了解，我们再次回到傅里叶级数。

傅里叶级数是用 $\cos n\varphi$，$\sin n\varphi$ 这样的项构成的，这些项在圆周上的情况，特别是它们的零点分布情况，是可以立即看清楚的。零点数是 $2n$，因而随着 n 加大而增加，此外，零点在圆周上是均匀分布的(图 26.11)。

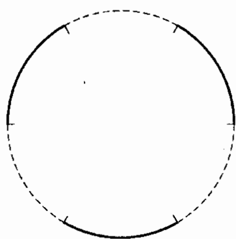


图 26.11

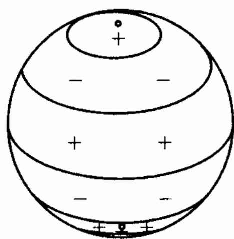


图 26.12

与此类似，首先考察函数 $P_n(\cos\theta)$，即 $P_n(z)$。

试问函数 $P_n(z)$ 在球面的零点情况如何？我们指出下面的情况，但不加证明，因为我们对球函数只是要获得一个大致概念。

$P_n(z)$ 在 $z=-1$ 和 $z=+1$ 之间有 n 个实零点，它们的位置对于 $z=0$ 对称；这就是说，在 n 个对于赤道对称的纬线上，$P_n(z)=0$。所

以，这个球函数 $P_n(z)$ 把球面划分为 $(n+1)$ 个球面带，在它们上面，$P_n(z)$ 交错地有正值和负值。因此，$P_n(z)$ 叫做“带形球函数”，这是英国先行者给予它的名称（图 26.12）。

在 $P_n(\cos\theta)$ 之外，我们还要考虑因子 $\left.\begin{matrix}\cos\nu\varphi\\ \sin\nu\varphi\end{matrix}\right\}\sin^\nu\theta$ 的作用。首先考虑最高次的款 $\nu=n$，这时 $P_n^{(n)}(\cos\theta)=1$，对应于球函数

$$\left.\begin{matrix}\cos n\varphi\cdot\sin^n\theta\\ \sin n\varphi\cdot\sin^n\theta\end{matrix}\right\}1。$$

关于它们的零点，主要涉及第一个因子（第二个因子给出球面的两极）。我们得

球函数 $\left.\begin{matrix}\cos n\varphi\\ \sin n\varphi\end{matrix}\right\}\sin^n\theta$ 把球面划分为 $2n$ 个以经线为边界的、角度相等的扇形（球面二角形），在它们上面，球函数交错地有正值和负值。这个函数因而叫做“扇形球函数”（图 26.13）。

现在考虑中间诸款：

$$\left.\begin{matrix}\cos\nu\varphi\ \sin^\nu\theta\\ \sin\nu\varphi\ \sin^\nu\theta\end{matrix}\right\}P_n^{(\nu)}(\cos\theta)。$$

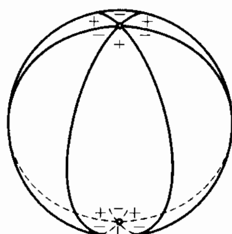


图 26.13

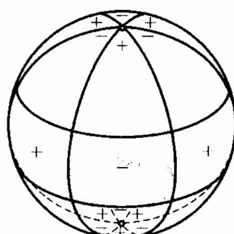


图 26.14

当第一个因子等于 0 时，得出 ν 条经线，它们把赤道分为等长的弧。第二个因子 $P_n^{(\nu)}(\cos\theta)$ 等于 0 时（不难看出）给出 $(n-\nu)$ 条对于赤道

对称的纬线。这样，球面就划分为“四角形区域”，只有和两极相接的区域例外，这些是三角形。因此，这些函数就叫做“四角形球函数”[①][②]。（图 26.14）

因此，正如 $\cos n\varphi$，$\sin n\varphi$ 可以称为在圆周上的振荡函数那样，我们的球函数对于球面也可用这个名称，叫做“球面振荡函数”，它在球面划分的域上交错有正负值。对此进一步作具体数值考察当然是特别有意义的，特别是可以看看不同球函数的正负值域是怎样相互覆盖的。

§26.5 用有尽球函数级数作近似表示的误差估计

这些振荡球函数构成我们所要讨论的级数问题是，它和函数 $f(\theta, \varphi)$ 在各点的值接近到什么程度，即我们能否估计其误差。

对此，我只回答如下：

这项误差，即有尽级数和它所要代表的函数之差，在每一点都可以用一个在球面的积分表达，完全对应于我们所知道的傅里叶级数的款。

这项估计见于教科书中；不过在那里，重点往往只是证明，当 $f(x, y)$ 是“合理”函数时，其余项随着 n 的增加而趋于 0；可是在实践中需要对于有尽的 n，对余项作出估计[③]。

在具体实践中，要把球函数展开式取多少项才能达到所需的准确度，那不是我们的理论研讨中的课题，需要根据不同情况分别加以

---

① [来源于希腊字“四”。]

② 四角形球函数的德文是 tesserale Kugelfunktion，希腊文的“四”是 Tesseres。——中译者

③ [无尽球函数的级数展开在许多方面和无尽傅里叶级数类似。参看 L·费耶尔的 “Über die Laplacesche Reihe”（“关于拉普拉斯级数”），*Math. Ann.* 第 67 卷，1909 年，第 76—109 页以及外尔的两篇关于球函数的吉布斯现象的文章，发表在 *Rendiconti del lircolo Matematico di Palermo*，第 24 卷（1910 年），第 308—323 页和第 30 卷（1910 年）第 337—407 页。]

判断。

一个著名的例子是 1839 年高斯的工作：Allgemeine Theorie des Erdmagnetismus（Resultate aus den Beobachtungen des magnetischen Vereins im Jahre，1838）[《地磁学的一般理论（1838 年磁学协会观察结果）》二文集，第 5 卷，第 121—193 页]。在那里，计算了直到四次球函数表示的地磁力的势能 V，其偏导数

$$\frac{\partial V}{\partial x}=X,\ \frac{\partial V}{\partial y}=Y,\ \frac{\partial V}{\partial z}=Z$$

给出地磁分力（只涉及 24 个系数，因为对于势能，作为常数项出现的那个系数不起作用。）。1839 年以后，根据更好的观测材料，对这个计算结果加以改进了，其中特别是诺伊迈尔（Neumayer）在汉堡的工作[①]。在那里，指出了用球函数作近似表示时，到第四次项就可以获得足够好的全貌，用到第五次项也不能对此改善。在这里，和在别处一样，高斯对应用的需要有着高度的敏锐感。其具体推算过程也是富有启发性的。

---

① 其合作者中特别要提到施密特[较准确的文献见施密特关于地磁学的报告，*Enzykl. der math. Wiss*，Ⅵ，第 10 卷，第 20 节。]

第八部分

平面曲线的自由几何

我们现在进入讲演的第二部分，我用的标题是“平面曲线的自由几何”。加上“自由”这个词是表示所讨论的定义及其推论与所选的固定直角坐标无关，这和至今为止讨论函数概念时的一般做法是不同的。为了节省篇幅，我们限于论述平面曲线。

这里也有精确数学和近似数学的分野；和以前一样，我们将要讨论的是解析几何的理想构造，它是（在公理的基础上）以实数概念为支柱的。和解析几何相对立的是综合几何，它也以相同的公理为基础，但又和图形自身打交道。我们采用解析的方法，不是由于本质的需要。下面涉及的所有问题都可以用综合几何方法处理，但是一方面若那样做，在联系到数学其他领域时，综合几何的储备还不足，因而会有所遗漏；另一方面，有关研究文献所用方法又几乎完全是解析的。

在这里，分析和几何的关系原则上和第一篇相同：利用分析把几何精确化，通过几何图像使分析更具活力。

第二十七章

从精确理论观点讨论平面几何

§27.1 关于点集的若干定理

我们从精确几何观点下的考察开始。我们将以点集论为基础来处理我们的对象，先把今后用到的最简单的定义综合列举如下：

（1）点的坐标(x, y)是近代实数概念意义上的实数，即它们是用有尽或无尽十进制数界定的，并具有正号或负号。

（2）无穷点集本身就是数学深入研究的对象。它们有着值得注意的各种不同特性[①]：

（a）一个点集可以可数或不可数。当它的元素可以通过任何一种方式和自然数建立一一对应关系时，它就称为可数的。

（b）一个节是指一个其边平行于坐标轴的矩形内部点的整体。特殊地若需要强调矩形的边界点不属于点集，它就称为开节。相反，若边界点也都属于点集，它就叫做闭节。平面上任意被包围着的一块中的一切点，称为闭区域或开区域（＝域）[参看第 102 页[②]]。

（c）若一个无穷点集(x, y)中一切点的坐标的绝对值都小于一

---

① [对于下文，最重要的文献是：1. 已经举出过的《百科全书》（ⅡC9）上 A・罗森塔尔的论述；2. C. Jordan, *Cours d'Analyse*（若当：《分析教程》）第三版，第一卷，巴黎，1909年；3. F. Hausdorff, *Grundzuge der Mengenlehre*（豪斯多夫：《集合论基础》）第二版，莱比锡，1927年；4. C・卡拉西奥多里：*Vorlesungen über reelle Funktionen*（《实函数讲义》）第二版，莱比锡与柏林，1927年；5. H. Hahn, *Theorie der reellen Funktionen*（H・哈恩：《实变函数论》）第一卷，柏林，1921年。]

② [关于区域概念的更准确讨论，参看第 140—142 页。]

个固定正数，点集就称为有界的。

(d) 已给一个无穷点集和一点 $P(x, y)$。若 xy 平面上每一个含 P 作为内点的节(或区域)也含有点集的无穷多点，则 P 称为点集的聚点。下面的定理成立：

每一个有界无穷点集 (x, y) 至少有一个聚点。

(e) 点集的一个聚点可以属于也可以不属于该点集。若点集含有它的一切聚点，它就称为闭集。

(f) 若点集的每个点都是它的聚点，它就称为自稠密(简称自稠)的。

(g) 自密的闭集称为完备集①。

(h) 已给 xy 平面上一个节(区域) I 里的一个点集，若 I 的每个子节(子区域)都含有该点集无穷多点，则点集称为在 I 里处处稠密。例如 xy 平面的有理点的集合在平面上处处稠密。若点集在 I 的任何子节(子区域)里都不处处稠密，它就称为在 I 里无处稠密。

你们看，这些都是容易理解的概念，但只有在具体事例中运用它们，人们才会领会其本质和意义。

§27.2 通过对两个或多个不相交圆的反演所产生的点集

我希望你们从头就确信，点集的研究对象是纯几何的问题。因此，我要举的例子，不是通常遇到的那种例子，那些例子表面上是用算术方式建立的，因而给予人们印象是人为的。我想采用一个点集的几何产生原则，这个原则是从自守函数理论发展起来的。

我们先考虑半径倒数变换或对(于)圆的反演②。

---

① perfekt. ——中译者

② 中文一般用“反演”。——中译者

设已给定一个半径为 ρ 的圆以及在圆外的一点 p。通过公式 $rr'=\rho^2$ 所表示的定则，由 p 得出圆内一点 p'，称为 p 的反演点（图 27.1）①。这种变换有时也叫做"对于圆的反射"。但这只能有保留地②来理解，因为它根本与光学上凸镜的反射规律无关，只是当圆退化为直线时，才出现光学意义上的反射。

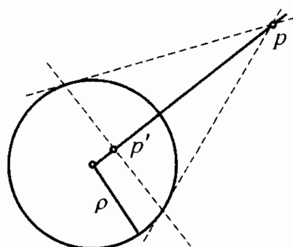


图 27.1

我们假定已经了解这个变换的一些性质：圆总是变成圆，因而两圆交角经过反演不变，但正向颠倒。在这里，要强调的是，直线也包括在圆的范畴内，它们是看做半径无穷大的圆。无穷远处表现为一个点，直线是经过无穷远点∞的圆。p 和 p' 的关系也可界定为：它们是和原来的圆正交的所有圆（它们构成一个圆束）的两个公共点。这个定义的优点是，它只涉及对任何其他反演不变的元素。

在自守函数理论中，将要把任意多个反演相结合。我们举下面的例子。设有 3 个圆 K_1，K_2，K_3，它们每个都位于其他两个之外。设 p 为位于 3 个圆外部的一点。我们作 p 的反演点 p_1，p_2，p_3，再从这 3 个点得到它们对 3 个圆的反演点，如此下去。这样所得到的点，用记号表示为

$$(p)S_1^{\alpha_1}S_2^{\alpha_2}S_3^{\alpha_3}S_1^{\beta_1}S_2^{\beta_2}S_3^{\beta_3}\cdots$$

其中 S_1，S_2，S_3 依次表示对 K_1，K_2，K_3 的反演变换，α_ν，β_ν，…表示整数，代表着运用各该反演的次数。由于每个反演都是对合的③，

① r，r' 依次表示 p，p' 和圆心的距离，p，p' 与圆心共线，并在圆心同侧，若 p 在圆上，则 $p'=p$。——中译者

② 原书用了拉丁文 cum grano salis，直译是"带着一颗盐"，意译是"有保留地"或"略带勉强地"。——中译者

③ 即每个反演的逆就是它自己，或者说，每个反演的"平方"是幺变换（或称恒等变换）。——中译者

S_1^2, S_2^2, S_3^2 都表示幺变换，α_ν, β_ν, … 的值可以限于 0 和 1。问题是，我们如何得到这些点所构成的集合的图像？这个点集有哪些聚点？

下面要考察的就是从一个已给点 p 出发，经过对 3 个圆作任何组合的相继反演所得到的点的集合。所以，我们面对的是整个“变换”群，它有 3 个“生成元”。所得到的点都简称为 p 的等价点。

若有人提出异议，认为这个问题虽然是纯几何性质的，可终究是人为的，那么可以指出，这里所论及的点集早在 1850 年就为汤姆逊和黎曼在静电学问题中考虑过了。当人们考察具平行轴的 3 个回转柱面（它们的正交横截面就是我们的圆）上电荷的平衡状态时，“电像法”①就导致我们点集的研究。

我们先从较简单款着手：设已给两个圆 K_1, K_2，每个都在另一个之外（参见图 27.2），其对应的反演用 S_1, S_2 表示。

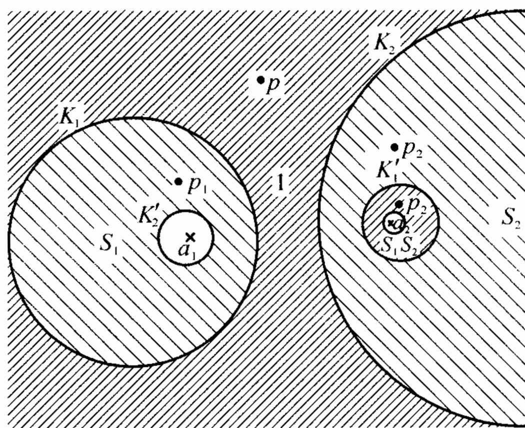


图 27.2

① Methode der elektrischen Bilder. ——中译者

首先改变作法，不是由单独一点 p 出发，而是取对两圆来说都属于外部的那个区[①]的一切点。这个区用 1 表示，并对它相继施行反演 S_1 和 S_2。这样得到的区称为和 1（因而彼此）等价。我们将看到。这些区一个接一个，相接处没有间隙，也不交叠[②]，直至覆盖除两个点 a_1 和 a_2 以外的整个平面，这两点就叫做极限点。只要对这些彼此等价的区有了整体的理解，就会得到对彼此等价点所构成的集合的整体理解。因为在最早的区中每个点 p，在每个等价区里都有自己的一个而且只有一个等价点。

对区 1 施行反演 S_1 和 S_2 所得的最早两个新区也用 S_1 和 S_2 表示。S_1 沿 K_1 无间隙地相连，但在其内部，它还有另一个圆作为边界，在图中用 K_2' 表示，因为它是（通过 S_1 得到的）K_2 的像。同样，S_2 也和 1 连接，它内部也有一个边界圆 K_1'。

现在要作出其他和 1 等价的区，以便得到这种区的全部。为此，我们首先让 S_1 和 S_2 分别对其内部的边界圆（即 S_1 对 K_2'，S_2 对 K_1'）作反演。我要证明，所得的新区，即 S_1 和 S_2 所提供的，进入它们里面的圆孔的那两个区，可以由 1 通过反演 S_1 和 S_2 的简单结合得到，因而属于所求的等价区之列。

证明如下：考虑对 K_1' 的反演。对于 K_1' 互为反演像的一对点，经过反演 S_2，变为对 K_1 互为反演像的一对点。这是因为，对于 K_1' 互为反演像的一对点是同 K_1' 正交的一切圆（这些圆构成一个圆束）的两个公共点，这个圆束，经过反演 S_2 变成同 K_1 正交的圆束，前一个圆束的公共点变成后一个圆束的公共点，即对于圆 K_1 的互为反演的一对点。所以，为了得到 S_2 对于 K_1' 的反演像，可以让区 1 先对 K_1 作反演得到 S_1，再让 S_1 对 K_2 作反演。于是那个新区可以用 S_1S_2 表示，因而的确属于和 1 等价的区之列。同样，S_1 对于 K_2' 的反

---

① 按定义，此处应作“域”（不含边界点），但原文作 Bereich（区域或区）；而按下文用法，译文以作“区”为宜。——中译者

② 指没有公共内点。——中译者

演可以用 S_2S_1 表示。

上述作图法引出"进展原则"[①],用它可以得到一系列和 1 等价的新区。我们将让每个新区(首先是区 S_1S_2 和 S_2S_1)再对其内部边界圆作反演。这样,一方面(图 27.2 右)得到一个无间隙地相连接的区序列 S_2, S_1S_2, $S_2S_1S_2$, $S_1S_2S_1S_2$, …;另一方面(图 27.2 左)得到相应的区序列 S_1, S_2S_1, $S_1S_2S_1$, …,它们分别无限制地缩小到上面指出的两个极限点 a_1 和 a_2。

事情美妙之处在于,按我们的进展原则所获得的区已经包括全部所求的互相等价的区。这是因为,由于 $S_1^2=1$, $S_2^2=1$,每一个由 S_1 和 S_2 两个反演变换所构成的乘积都可以化为 S_1 和 S_2 交错结合的形状。因而是按我们进展原则所产生的区之一,若这个系列的反演中最后一个是 S_2,它就出现在图 27.2 右方,否则就出现在图 27.2 左方。

这样,通过我们的进展原则,就得到所有互相等价的区,从而也就得到关于 p 的一切等价点的概貌。我们同时指出,这整个图形不但对于 K_1 和 K_2 是自己的反演像,它对于每两个区的共同边界图,例如对于 K_1' 或 K_2' 也是。试对 K_1' 加以证明如下:我们已经知道,通过对 K_1' 的反演,区 S_2 变成区 S_1S_2。所以,对 K_1' 的反演可以用 $S_2^{-1}S_1S_2$ 表示,而由于 $S_2^2=1$,即可用 $S_2S_1S_2$ 表示。现在,若令这个变换作用于任意一个等价区(它本身就可以用 S_1,S_2 相继结合构成),只需把 $S_2S_1S_2$ 续在后面即可,这样,所得到的反正是另一个等价区。这就完成了证明。证明了这个定理,也同时得到我们进展原则的一种推广。

迄今为止,我们只是用 K_1 作为 S_2 的边界来作反演,以由 S_2 得到那个也以它为边界的区 S_1S_2。显然,我们也可以把早已得到的 3 个一个接一个的区 S_2,1,S_1 同时对 K_1' 作反演,一下子得到 3 个新区,即 S_1S_2, $S_2S_1S_2$, $S_1S_2S_1S_2$。这样,就一共有 6 个一个接一个的区,这 6 个区又可以对 $S_1S_2S_1S_2$ 的内部边界圆作反演,一下子又得

① Fortsetzungsprinzip, Fortgangsprinzip。——中译者

到 6 个新区，如此等等，这个步骤就称为扩大了的进展原则。

上述从一个图形逐步跟踪下去的情况提供一种方法来处理更困难的款，那就是，已给多于两个圆，把对它们的反演无限制地相继结合起来。我们首先考察最早提出的 3 个互不包含的圆 K_1，K_2，K_3。

我们考虑 K_3 为直线的特款(参看图 27.3)。为了得到以圆 K_1，K_2 和直线 K_3 作为共同边界的区 1，显然 K_1，K_2 必须或者都在 K_3 的左侧，或者都在右侧。我们让区 1 对 K_1，K_2，K_3 作反演以得到区 S_1，S_2，S_3，它们内部各有一个边界圆 K_2'，K_3'，如此等等。进展原则要求把每个新区 S_1，S_2，S_3 分别对其内部边界圆再作反演。每次又得到一个区，其内部各有两个边界圆。对这些内部边界圆再次作反演，如此无尽地继续下去。这样，整个平面就被无穷多个彼此等价的区"单一地"覆盖①，同时也被无穷多个彼此等价点所构成的那些集合单一地覆盖，只有极限点是例外。这些极限点位于所得到的各圆孔内，这些圆孔又无限制地缩小到极限点，而其数目却又无限制地增加。现在有一个问题：在较少的反演之后，能否对极限点所构成的集合作出合理的预测？

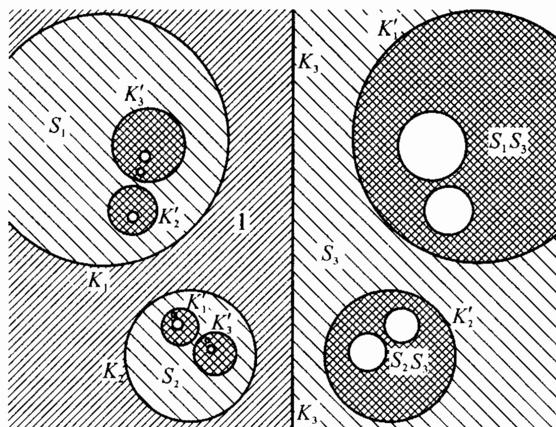


图 27.3

① 各边界圆上的点属于两个相邻的区。——中译者

我们先把主要情况综合为 3 点：

(1) 我们不是对单个点 p，而是对整个区 1(基本区)施行变换。

(2) 从最早的区以得到新区的过程，就是把新区以及已得到的全部图形再对新出现的边界图作反演(原来的和扩大了的进展原则)。

(3) 整个平面被最早的区以及它的无穷多个像无间隙地覆盖，只有极限点是例外，它们是在我们的步骤继续进行中出现的。

我们当然可以考虑从 4 个、5 个或更多个圆出发。那时将出现一些复杂情况，下面即将说明。不过，我要强调，3 个圆的图形还有一个简单特点。

我们知道，3 个圆有唯一的共同正交圆。对于反演 S_1，S_2，S_3，这个正交圆变为自己。因此，它也被这些反演的任何结合变为自己。由此可见，一切由这 3 个最早的圆经过反演所得的无穷多个圆也都和这个不变圆正交，因而由基本区产生的一切区都和不变圆相交，并且经过对不变圆的反演不变。这些区沿不变圆彼此相接地排列。所以一切极限点就都在不变圆的圆周上。这是很令人满意的，因为这样就使我们对极限点位置的判断告一段落。

当开始有 4 个或更多个互不相交的圆，而这些圆构成一个区的共同边界时，在特殊情况下，这些圆也会有一个共同的正交圆，这时极限圆情况就像上面所说的那样。可是一般地，共同正交圆并不存在。这样，就没有一条阿里阿兹内①的线来可靠地把我们导出等价区的迷宫，并引到极限点。

§27.3 极限点集的性质

可是我们仍愿意对一般款中的极限点集作出些结论。首先我重

---

① 希腊神话载：明诺斯王的女儿阿里阿兹内把一根线给了塞色乌斯，使他找到走出迷宫的途径。——中译者

复一下，在圆数 $n=3$ 的款，已无法确定这个点集的具体几何状况，更不用说 $n>3$ 的款了；我们的表现能力在最早几次反演中就要显出局限性[①]；而且在这里，关于无穷，席勒(Schiller)的诗句(《世界的伟大》)也适用：

收敛起遐想的翅膀，

勇敢的航海者

就此失望地抛下锚罢。

但是令人赞赏的是，利用点集的概念性定义，人们还的确能作出某些结论。这样做之所以可能，应归功于康托，是他的特殊贡献第一次使无穷成为数学研究的对象。

设画在页旁的圆(图 27.4)是我们映射中一个边界圆，对它反演，就把在它的外部的一切细节都映射到其内部。显然，它的内部不会被极限点填满。相反，那里首先有着不含极限点的区，又有新的边界圆。由此可见，不存在一个有穷区域或一段曲线，其中(或其上)极限点会处处稠密。于是有第一个结论：

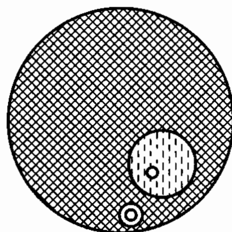


图 27.4

(1) 由我们的进展原则可知，极限点集无处稠密。

现在取一个极限点，以及包围着它的一个小边界圆(图 27.5)。由于外部整个图形要经过反演变到这个边界圆之内，其中也包括无穷多个极限点，可见这些极限点可以任意地靠近原先所取的那个极限点。换句话说，我们图中每个极限点本身是无穷多个极限点的聚点。

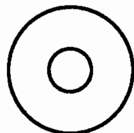


图 27.5

因此，极限点集是自稠的。

我们还要考察，这个点集是否甚至完备，即它是否含有它的一切

① 这是一个特别美妙的例子，用来说明我的论断：精确数学的事物中蕴含着超出我们想象力的东西。

聚点。设 p 为极限点的一个聚点，则 p 的每个无论多小的邻域都含有无穷多个极限点。如果假定 p 本身不是极限点，我们将立即发现矛盾。这是因为，若 p 不是极限点，它就或者是某个等价区 B 的一个内点，或者是某个等价区 B 的内边界圆上一点。在前一种情况下，p 就有一个只含 B 内点的邻域，那里面不会有极限点，同假设矛盾。在后一种情况下，我们把 B 对其内边界圆作反演，这时由 p 和 B 的反演像可知，p 有一个邻域，其中只有 B 的内点和 B 的边界点，因而没有极限点，这仍然同假设矛盾。于是得到定理：

(2) 极限点集是完备的。

按照我们根据有穷概念所形成的习惯，定理(1)和定理(2)显得不协调。而在这里，涉及无穷点集时，它们却是可以相容的。

可以进一步问，极限点的势①是什么？每个极限点完全确定于一个由数字 1，2，3 所构成的无穷序列，不过在这个序列中，两个相连数字不相同。

现在的问题是，这样的无穷序列有多少？

由 3 个数字这样构成的序列的势，和二进制分数(即按 $\frac{1}{2}$ 的幂展开的分数)的势一般大。对这个事实，我不作进一步阐明。于是我们所得到的是连续统的势。

你们看，与很简单的自守图形相联系，有着点集论中最美妙的性质。

自守函数理论引出许多比上述更复杂的图形，参看(*Fricke-Klein: Theorie der automorphen Funktionen*(《自守函数理论》)，第1卷，莱比锡，1892年，第428页等中的大量图形)。一般地，在几何中只要把一种作图步骤按一定规律无尽多次地连接进行，就必然要遇到无穷多个点或无穷多个域的集合。这表明，点集论不仅对于自守函数理论的几何图形有着根本意义，它对几何的其他部

① Mächtigkeit(英文 Power，Potency，Cardinal number). ——中译者

分也同样是如此。

其次，这一切当然是精确数学。在近似数学中这里所说的完全失去其意义。

第三，我愿意表达一种希望：点集论固然对几何提供助力，它也应从几何获得新的动力。点集论专家们所推出的例子往往显得带有人为成分，可是，像我们所看到的，当我们从几何出发时，却自然地会引出纯粹属于点集论的问题。在这里，几何有着最令人惊异的推动能力①！

---

① ［近来点集论对于几何分支中的拓扑学，在奠定其基础及推动其发展中，都起了决定性作用。此外，它在可展曲面、小积曲面、有尽连续群理论的研究中，以及在矢场和凸图形的研究中都有应用。与此有关的文献可参看已多次举出的、罗森塔尔在百科全书里的文章中的第 1013 页。

要对第 140—154 页所介绍的关于维数精确概念的发现作补充，似乎是不可能的，那里介绍的概念是点集拓扑学中最新成就。维数的新定义比原有的定义远为广泛，却又结合着朴素的直观。它是最近独立地而且几乎同时地为维也纳的卡尔·门格尔（Karl Menger)和莫斯科的 P·乌里松(Paul Uryson，1898—1924 年)发现的。不过 H·庞加莱在其短文“Pourquoi l'espace a trois dimensions”(“为什么三维空间”，(*Revue de Métaphysique et de Morale*，第 20 卷，1912 年)，第 484 页，在 *Dernières Pensées*(《最新思想》)中重印)中已经指出了解答这个 2 000 年老问题的途径。至于这个老问题，则在欧几里得几何原本中一开始就出现了。新的维数概念发展成一个广泛的理论，但不久前它的主要结果才第一次部分地有详细的阐述，那里 P·乌里松有篇伟大的 *Mémoire sur les Multiplicités Cantorien nes*(《关于康托的众多工作的报告》，*Fundamenta Mathematicae* 第 7 卷，1925 年和第 8 卷，1926 年；续篇在 *Verhandelingen der Amsterdamer Akademie*，1927 年)，还有 K·门格尔的好几篇在 *Wiener Monatsheften für Mathematik und Physik* 以及 *Amsterdamer Akademie* 上发表的文章。进一步的介绍可参考 *Jahresberichtes der Deutschen Mathematikervereinigung*，第 36 卷的底注(斜页码第 9 页起)所列有关报告的粗略说明。此外，关于已展开的思想内涵的简短概述，可以看门格尔本人所著 *Bericht über die Dimensionstheorie* (《关于维数理论的报告》，*Jahresbericht der Deutschen Mathematikervereinigung*，第 35 卷，1926 年，第 113—150 页)以及 P. Alexandroff 的两篇著作：*Darstellung der Grundzüge der Urysohnschen Dimensionstheorie*(亚历山大罗夫：《关于乌里松维数理论基本概念的阐述》)，*Math. Ann.* 第 98 卷，1927 年，第 31—63 页和 W. Hurewicz，*Grundriß der Mengerschen Dimensionstheorie*(胡维尔兹：《门格尔维数理论概要》，同上，第 98 卷，1927 年)，第 64—88 页。上面所提到的定义的基础是邻域概念(参考 F·豪斯多夫：《集合论基础》第二版，第 288 页)，它是下述直观上确凿事实的逻辑抽象：当我们取一点连同它的一个小邻域时，我们由三维体分离出二维曲面，由二维曲面分离出一维曲线，由一维曲线分离出由两点构成的零维点集。维数新定义当然是这样的：经典意义上的 n 维图形(例如 n 维欧氏空间的 n 维立方体或 n 维空间本身)按新定义也是 n 维的。证明见门格尔和乌里松的文章。］

§ 27.4 二维连续统概念、一般曲线概念

紧接着的下一个问题也是在函数论的土壤上生长出来的，它最早为魏尔斯特拉斯在他的讲义中给了必要的答案。问题是：

什么时候我们把一个二维点集（即平面上的一个点集）称为连续统（或区域）？

在平面点集中，我们称为连续统的最简单的例子是一个圆（或长方形）内部的一切点的集合；在这里，我们约定，边界点不包括在内。

在平面连续统点集的一般定义中，我取出以下两项性质[①]：

(1) 首先要求集合是“连通的”，即集里任意两点总可以用一条有尽多个边构成的折线相连，折线的点都属于集。

(2) 其次，以集里每点为中心，可作一圆，圆内部的点也都属于点集。

现在，关于这样一个点集的边界问题，按照我们已有的知识，已经可以举出各种各样的例子。

(1) 在平面上，除去一点，所有其他点都可以属于点集。

(2) 回顾第 135—139 页所论的自守函数的例子，我们看到，一个连续统的边界可以构成一个无处稠密的无穷集。

(3) 边界还可以构成一条普通意义上的曲线（圆、长方形等等）。

(4) 但还有别的边界。下面是 W · F · 奥斯古德（Osgood）所考虑过的[②]。

例如，只考虑正半平面（x 轴上方的一切点）；并且在 x 轴上某些处作垂直于它的直割线（图 27.6）[③]，和半平面上别的点不同，这些割

---

① ［根据新近拓扑学中康托的定义，连续统是含不止一点的连通闭集。书中界定的“平面连续统”概念等价于欧氏平面的一个连通域。］

② 例如参看 *Transactions of the Amer. Math. Soc.* 第 1 卷，1900 年，第 310—311 页。

③ 这些割线的长是有穷的。——中译者

线上的点都不看做是区域的点。现在设想，让这些割线不断增加，若它们和 x 轴任意一段上的交点都不处处稠密，则割线之间总有区域的带伸到 x 轴。x 轴上这样的无处稠密、完备、具有连续统的点集存在，

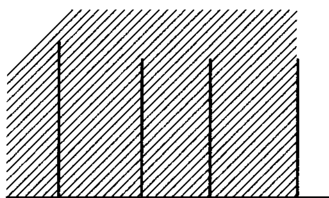


图 27.6

前面已看到了(回顾一下 3 个圆的款的自守圆形中的正交圆上的极限点)。由此可见，一个连续统可以有无穷多条割线，其排列位置就像直线上一个无处稠密、无穷而完备的点集。

这 4 个例子表明，一个区域可以有多么多样化的边界(人们往往不假思索地以为它是一条曲线)，因此，有必要谈谈在当前数学中关于这个问题的观点①。

一个连续统的边界可能有各种各样的不同情况，因而贸然地把曲线界定为一个区域的边界是不恰当的。当然，这是对精确数学领域说的。而精确数学是理想化地建立在现代实数概念的基础上的，特殊地，理想化图形又是从空间概念引出来的。那么，一条曲线的定义应当是怎样的呢？我们作如下说明：

令一个变数 t 经过一个闭节 $a \leqslant t \leqslant b$ 中的一切值，并令一点的坐标等于这个参数 t 的单值连续函数 $x=\varphi(t)$，$y=\psi(t)$，这样界定的点就构成一条曲线。用纯语言表达，曲线的定义是：

一条平面曲线是平面上一个点集，它是直线上一个闭节的单值连续映像。

通过参数 t，人们不自觉地把曲线这个定义和力学观点混同。人们可以把 t 看做时间(这也是习惯上用字母 t 的根源)，并说，当参数 t 从时刻 a 到时刻 b 时，点 $x=\varphi(t)$，$y=\psi(t)$ 描绘曲线。换句话

① [在这里和下文可参看 A・罗森塔尔："Über den Begriff der Kurve"("关于曲线概念")，*Unterrichtsblätter für Mathematik u. Naturwissenschaften*，第 30 卷，1924 年，第 75—79 页。还可参阅第 150 页底注①。]

说，曲线是在一段时间里一个点连续运动的轨迹。这一切都是很容易懂的。较困难的问题是，这样界定的一条曲线有什么样可能的表现？

§27.5 覆盖整个正方形的皮亚诺曲线

在这里，我必须提请注意 G·皮亚诺的一项发现，它发表在 1890 年①，于 1891 年由希尔伯特在几何上加以阐明②。它指出，用单值连续函数 $x=\varphi(t)$，$y=\psi(t)$界定的一条曲线可以完全覆盖一块平面区域。这样的一条曲线叫做皮亚诺曲线，不难给出这样曲线的实例。

我先要说明，这样一条曲线，不是指像下面所说的那种曲线。因为有人听到皮亚诺的成果时，可能会认为它是和外摆线有关的一桩旧事物。

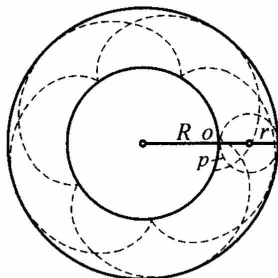


图 27.7

设在一个半径为 R 的圆的外侧，有一个半径为 r 的圆在滚动，而比例$\frac{r}{R}$是无理数。在滚动的圆上任意选定一点，它就描画出一条不封闭的外摆线，因而它所在的环形区域，包括边界在内，被处处稠密地覆盖了（图 27.7）。但这并非表明，环形域的每一点都在外摆线上，而只是说，给定环形域的任意一点，只要沿着外摆线不断地向前走，就能走到离该点任意近的地方。在区域边界上和内部，都不难给出曲线达不到的点。

假定在滚动开始时，产生曲线的点就是固定圆和动圆的切点，在图中我们用 o 表示。设想在固定圆上取一点 p，它和 o 之间的弧长

① *Math. Annalen*，第 36 卷，1890 年，第 157—160 页。
② *Math. Annalen*，第 38 卷，1891 年，第 459—460 页。

是$\frac{m}{n}2\pi R$,其中 m 和 n 是整数。由于$\frac{R}{r}$是无理数,长度$\frac{m}{n}2\pi R$ 和动圆的周长 $2\pi r$ 是无公度的,因而这样的点 p 都不在外摆线上。再取这样一个和固定圆同心的圆 k,它被外摆线两个相继的弧截出长度为 b 的弧,则 b 和圆 k 的周长 u 必无公度。否则容易看出,外摆线就将是封闭的。现在设 a 是外摆线在 k 上的一点,而 q 是 k 上另一点,它和 a 之间的弧长是$\frac{m}{n}u$,其中 m 和 n 仍是整数。从 a 出发无论多少次在 k 上截出弧长为 b 的一段,永远达不到 q。因此,点 q 不在外摆线上。

所以,不管这条外摆线多么有趣,决不是皮亚诺定理中所讨论的曲线,那里所说的曲线要对于一定的 t 达到某块区域里的每一点。

现在对皮亚诺曲线作几何说明。

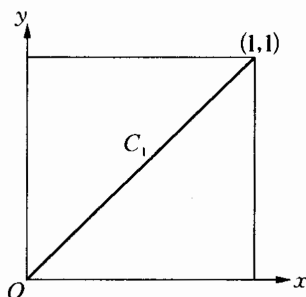


图 27.8

作为例子,我们选取把 xy 平面上的正方闭节 $0\leqslant x\leqslant 1$, $0\leqslant y\leqslant 1$ 完全覆盖的一条曲线。我们界定这条曲线为曲线序列 C_1, C_2, C_3,…的极限曲线 C_∞。在这里,设 C_1 是正方形从坐标原点(0, 0)到点(1, 1)的对角线(图 27.8)。我们采用 C_1 从(0, 0)量起的"弧长"除以$\sqrt{2}$,作为参数 t。这样,函数 $\varphi_1(t)$和 $\psi_1(t)$就是 $x=t$, $y=t$,其中 $0\leqslant t\leqslant 1$(图 27.9)。

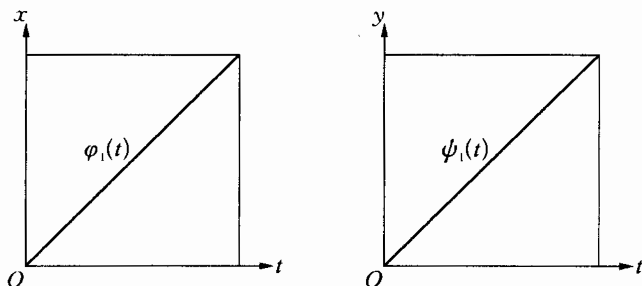


图 27.9

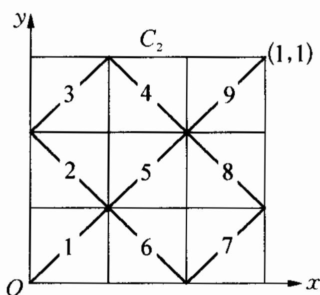


图 27.10

现在到了决定性的一步，要按这一步作出 C_2，C_2 的端点依然是原来正方形对角线的始点和终点。但我们把该正方形分为较小的 9 个[①]，并把这 9 个小正方形中的对角线像图 27.10 那样连接起来，以代替大正方形的对角线，从而得到 C_2。这个图指出了 C_2 从一个小正方形到另一个的转折点，使 C_2 各部分的顺序较易了解。我们看到，大正方形内部有两个点，C_2 经过它们各两次，即小正方形 1，2，5，6 的公共点和 4，5，8，9 的公共点。新曲线 C_2 的长显然为 C_1 的 3 倍，即 $3\sqrt{2}$。取 C_2 的弧长除以 $3\sqrt{2}$，作为它的参数 t，就得到代表 C_2 的函数 $x=\varphi_3(t)$，$y=\psi_2(t)$，这两个函数的图就是图 27.11 和图 27.12 中的折线 $(0\leqslant t\leqslant 1)$。由此可见，$\varphi_2$ 的曲线可以如此得到：把 ψ_2 的曲线按比例 1∶3 缩小，再把所得 3 条折线沿对角线 φ_1 接起来。曲线 ψ_2 本身则可以如此得到：把原正方形的曲线 φ_1 横向按比例 1∶3 压缩，然后把所得线段及其对纵线的反射像交错地接起来。

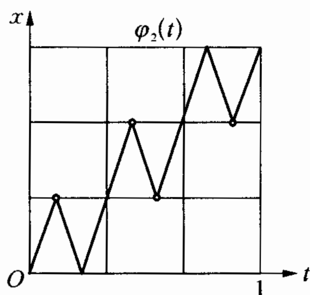


图 27.11

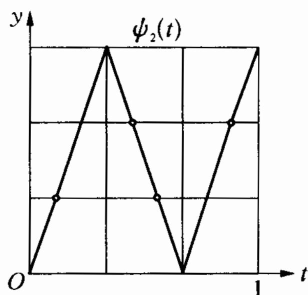


图 27.12

① 可以用大于 3 的其他奇数的平方代替 9，不需作本质改变。

现在我们把上述由 C_1 到 C_2 的步骤在 9 个小正方形分别重复一次，并如此无限制地继续下去。

就像通过把正方形分成 9 个来从 C_1 得到 C_2 那样，我们通过再次把每个小正方形分成更小的 9 个来从 C_2 得到 C_3，这时要把已得到的每条对角线代以一条折线，这折线或者与缩小了的 C_2 相同，或者是缩小了的 C_2 的反射。

类似地，我们作出 C_4，等等。

要画出 $y=\psi_3(t)$，我们把节 $0\leqslant t\leqslant 1$ 分成 9 段，同时也把节 $0\leqslant y\leqslant 1$ 分成 9 段，以得到曲线 $\psi_3(t)$，它含有 27 条线段，每一条由 3 小段(项)合并而成，故共有 81 项，对应于曲线 C_3 的 81 条小对角线[图 27.13①]。

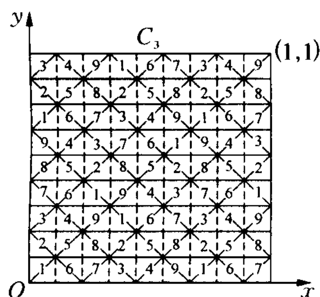


图 27.13

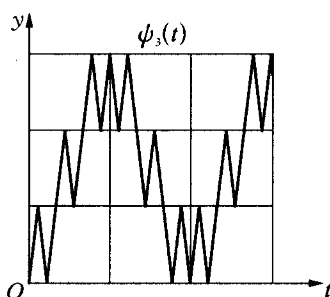


图 27.14

若把 ψ_3 和 φ_2 比较，就可以看出，由 ψ_2 得到 ψ_3，就像由 φ_1 得到 ψ_2 那样。

现在又可以从曲线 ψ_3，作出 φ_3 如下：把已经画出的 ψ_3 按 1∶3 的比例缩小，然后把这样所得的图沿 φ_1 的对角线相接。

但我们还可以由 φ_2 得到 φ_3，由 ψ_2 得到 ψ_3。一般地，有：

为了由 φ_n 得到 φ_{n+1}(或者，$n\neq 1$ 时，由 ψ_n 得到 ψ_{n+1})，我们先把

① 编辑将此处的图 27.14 改为图 27.13[在图 27.14，设想所画出的 9 个正方形都像图 27.13 那样再分。]

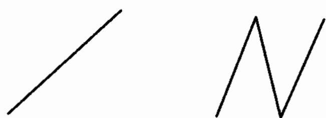


图 27.15

φ_n（或 ψ_n）的每个直边分成等长的 3 项，然后把每项代以一个（含 3 小段）的锯形钩，钩的始点和终点都和直线项相同（图 27.15）。

对曲线 φ_n 和 ψ_n 的这种前进规律，使我们想起魏尔斯特拉斯曲线的部分曲线，不过这里所遇到的不是正弦状的振荡而是直齿的锯形振荡。n 越大，锯齿的高度越下降，但其宽度则不成比例地急剧减少，因而越来越陡①。

我现在试着对极限曲线 φ_∞，ψ_∞，C_∞ 作进一步考察。

我们把 φ_n 和 ψ_n 的始点和终点，以及这两条曲线上两不同直线段相接处都叫做关节点（这样，ψ_3 和 φ_3 各有 82 个关节点），于是，这些关节点也是 φ_{n+1} 和 ψ_{n+1} 以及以后的一切曲线的关节点。在曲线 φ_∞ 和 ψ_∞ 上，这种关节点是处处稠密的。设想对每个 n 作曲线 φ_n 和 ψ_n 的关节点，并注意它们都属于曲线，则当 n 增大时，就得到连续曲线 φ_∞ 和 ψ_∞ 的全貌。

现在回到曲线 C_n，则除始点和终点外，φ_n 和 ψ_n 的关节点对应于 C_n 的转折点，在转折点处，两条对角线垂直地相接，而且构成 C_n 的多重点（如果只考虑原正方形内部的转折点的话），它们也是 C_∞ 的多重点。一条 C_{n_0} 的转折点，对于后面的 $C_n (n > n_0)$ 仍然是转折点，不过在两个转折点之间还不断添上新的转折点。

我们现在证明曲线 C_∞ 完全覆盖正方形 $0 \leqslant x \leqslant 1$，$0 \leqslant y \leqslant 1$。作 C_2 的图时，又把每个小正方形等分为 9 个更小的，如此等等。每个或大或小的正方形对应于线段 $0 \leqslant t \leqslant 1$ 上一个子节。现在，在原正方形边上或内部的任意点 Q 可以看做一个无尽序列的、一个套一个的正方形的极限点，这些正方形的对角线 J_n 随着 n 的增大而无限

① ［函数 φ_∞ 和 ψ_∞ 的连续性将在第 148—149 页证明。它们虽然连续，却不可微，因而是“魏尔斯特拉斯函数”。参看 L. Bieberbach, *Differentialrechnung*（比贝尔巴赫：《微分学》第二版，莱比锡，1922 年，第 107—111 页）。］

制地缩减到 0。由图可以看出，这套正方形对应于 t 轴上线段的节套，它确定节 $0 \leqslant t \leqslant 1$ 上一点 P。因此，由点 Q 至少可以得出一个参数值 t_0，因而 Q 必在 C_∞ 上。联系着我们的图，还不难证明，对应于每个 t 值，必有皮亚诺曲线上唯一的一点，因而由线段 $0 \leqslant t \leqslant 1$ 到正方形 $0 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant 1$ 的映像是单值的。为了证明这个命题，我们引进下面对子节的计数方法。

(1) 线段 $\overline{01}(0 \leqslant t \leqslant 1)$ 上子节的计数法。

第一次细分 T_1 所得子节，从左到右依次用

$$\delta_1,\ \delta_2,\ \cdots,\delta_9$$

表示。在第二次细分 T_2 中，每个 δ_ν 分为 9 等份，所得新子节，仍从左到右计数。这样，在 δ_1 里得子节

$$\delta_{11},\ \delta_{12},\ \cdots,\ \delta_{19},$$

在 δ_2 里，得子节

$$\delta_{21},\ \delta_{22},\ \cdots,\ \delta_{29},$$

$$\cdots\cdots$$

在 δ_9 里，得子节

$$\delta_{91},\ \delta_{92},\ \cdots,\ \delta_{99}。$$

我们设想，细分和计数都无限制地持续下去。

若线段 $\overline{01}$ 上一点 P 不是任何细分中子节的端点，就有唯一的一个节套确定 P。但若 P 是某个子节的端点，则（除 0 和 1 两点外）有两个节套都可以确定 P。其中一个节套的各子节都以 P 为左端点，因而从某个细分起，其下标止于 1；另一个节套中的子节，则以 P 为右端点，因而从某个细分起，其下标止于 9，例如 δ_{12} 的右端点，则由节套

$$\delta_1,\ \delta_{12},\ \delta_{129},\ \delta_{1\,299},\ \cdots$$

和

$$\delta_1,\ \delta_{13},\ \delta_{131},\ \delta_{1\,311},\cdots$$

确定。

(2) 正方形$\overline{01}$($0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$)上子节计数法。

由第一次细分所得 9 个小正方形按图 27.10 那样标上号码 1 至 9,记作

$$\eta_1, \eta_2, \cdots, \eta_9。$$

由第二次细分,从 η_ν 所得小正方形记以

$$\eta_{\nu 1}, \eta_{\nu 2}, \cdots, \eta_{\nu 9} (\nu = 1, 2, 3, \cdots, 9)。$$

在这 81 个小正方形的计数中,注意以下两点:

a) 没有一个正方形重复了或跳跃过去,先 η_1 后 η_2,以及以后的都加下标 1 到 9。

b) 由η_1 到 η_2,一般地,由 η_ν 到 $\eta_{\nu+1}$ 也没有跳跃,$\eta_{\nu 9}$ 和 $\eta_{\nu+1}$ 相差一个号码的则和它有一条公共边(图 27.13)。

设想细分和计数都无限制地继续下去。

我们看到,每个一维节 δ_{α_1},δ_{α_2},$\cdots$, δ_{α_ν} 对应于唯一的一个二维节 η_{α_1}, η_{α_2},$\cdots$,η_{α_ν},因而每一个节套(δ)对应于唯一的节套(η)。因此,若在$\overline{01}$上的点 P 不是某个子节的端点,它就对应于原正方形的一个内点 Q。当 P 位于 0 或 1 时,这当然也是对的。另一方面,若 P 是某个子节的端点,而(δ)和(δ')是确定它的两个节套,则根据计数步骤 b),其对应的二维节套(η)和(η')仍然确定唯一的 Q。于是已经证明了,每个参数值 t 只对应于皮亚诺曲线上唯一的一点。

最后,我们来证明函数 $x = \varphi_\infty(t)$ 和 $y = \psi_\infty(t)$ 的连续性,也就是皮亚诺曲线本身的连续性。设 P 和 P'是节 01 上两点,对应于参数值 t 和 t',(δ)和(δ')是确定它们的节套。若 t(或 t')是一个子节的端点,则总令(δ)[或(δ')]为以 t(或 t')为右端点的节套;我们预先排除 t 或 t'在 0 的款。假定节套(δ)和(δ')有 n 个公共子节。现在,若 Q 和 Q'为 t 和 t'在皮亚诺曲线上的像,(η)和(η')是确定 Q,Q'的二维节套,则(η)和(η')也有 n 个公共二维子节。令 t'无限制地靠近 t,则 n 无限制地增加,同时(δ)和(δ')的公共子节的最小长度无限制地

靠近零。由于当 n 增加时，二维子节的对角线无限制地缩短，就得

$$P' \to P \text{ 时}, QQ' \to 0。$$

这样就证明了 3 条曲线 $\varphi_\infty(t)$，$\psi_\infty(t)$，C_∞ 在节 $\overline{01}$ 的一切点连续。（当 p 在 0 点时的处理是简明的）

§27.6　较狭义的曲线概念：若当曲线

我们现在提出下面自然要有的问题：

皮亚诺曲线并不像生活中所遇到的曲线。如何对曲线 $x=\varphi(t)$，$y=\psi(t)$ 的定义加以限制，或者，如何要求这样界定的曲线的主要性质，使所得曲线和经验中的曲线更加类似？C·若当在他的分析教程中作了解答。

我们的皮亚诺曲线有无穷多个转折点，因而有无穷多个重点（进一步考察表明，它甚至有无穷多个三重点和四重点）；对应于多重点的 t 值在节 $0 \leqslant t \leqslant 1$ 上处处稠密。C·若当要求的是，公式 $x=\varphi(t)$，$y=\psi(t)$ 所界定的曲线在定义节内部没有多重点，即不存在两个或更多的 t 值 t_1，t_2，…（$a < t_1 < b$，$a < t_2 < b$，…），使

$$\varphi(t_1) = \varphi(t_2),\ \psi(t_1) = \psi(t_2),\ \cdots$$

对于节的端点 a 和 b 没有规定这个条件。若端点也满足这个条件，则 $x=\varphi(t)$，$y=\psi(t)$ 代表的图像称为开若当曲线。若 $\varphi(a)=\varphi(b)$，$\psi(a)=\psi(b)$，则曲线始点和终点相合，曲线称为闭若当曲线。因此，开若当曲线是一条线段的双向单值连续像，而闭若当曲线是一个圆的双向单值连续像。

有以下对于分析的基本定理：

每条闭若当曲线把平面分成两个以它为边界的连通域。

为了说明这个定理并非自明的，我指出，皮亚诺曲线之所以有异常表现，问题不出自事物本身，而出自名词的运用：因为在那里，我

们使用“曲线”这个名词时，其范畴比从经验曲线所获得的类似观念要宽广得多。为了把若当曲线所引出的问题尽可能讲清楚，我们先不谈若当曲线而谈若当点集，然后指出，涉及连通和分隔等问题时，若当点集和经验几何中的普通曲线是一致的。事实上，如果我们直截了当地说，每一条用 $x=\varphi(t)$，$y=\psi(t)$ 界定的闭曲线，若满足若当条件，就把平面分隔成一个内域和一个外域，听起来是浅显的。这里的根本问题在于，人们使用曲线这个词时无意中有着两种不同的含义。也许人们本应更详细地这样讲，由经验领域出发，认为每一条闭曲线把平面分隔为一个外域和一个内域，是简单明了的，但还必须问：在理想领域，对一个点集应如何进一步界定，类似的定理才能成立[①]？答案是：为此，它必须满足若当条件。这时，那样的点集才能叫做曲线！这样，定理本身的含义就说明了，至于证明，我只能介绍其主要观点，其细节最好还是阅读若当的分析教程。

设想已给满足若当条件的点集 $x=\varphi(t)$，$y=\psi(t)$。C·若当开始其证明如下：作一个无尽序列的闭多边形 P_1，P_2，P_3，…其中每一个把前一个包围在内，但都不含点集的点在其内部；再作另一个无

① ［第 139 页提到的维数概念对解决这个问题也提供一个合理的曲线定义：一条曲线是一个紧致的一维连续统。在这里，若一个连续统 K 的每个无穷子集都有一个属于 K 的聚点，K 就叫做紧致；由于连续统是闭集，聚点当然都属于 K。由朴素的直观，若对非紧致一维连续统添上∞点，所得也可以叫做曲线（如抛物线）。

这个曲线定义和直观一致，它不包括像皮亚诺曲线那样的点集。它是所谓康托曲线概念的推广。所谓康托曲线是欧氏平面上一个不含内点的有界连续统。关于曲线理论的文献可以举出：K·门格尔：“Grundzüge einer Theorie der Kurven”（“曲线的一个理论的基础”）*Math. Ann.* 第 95 卷（1925 年），第 277—306 页，以及在 *Zeitschrift*, *Amsterdamer Proceedings*, *Fundamenta Mathematicae* 上发表的论文。乌里松与此有关的研究在以下著作的第二部分可以找到：“Mémoire sur les Multiplicités Cantoriennes”（“关于康托的众多工作的报告”）（*Verhandelingen der Amsterdamer Akademie*, 1927）最后还可参考亚历山大罗夫：*Über kombinatorische Eigenschaften allgemeiner Kurven*（《关于一般曲线的组合性质》）。*Math. Ann.* 第 96 卷（1926 年），第 512—554 页。

在至少二维的欧氏空间里，每条曲线是一个域的边界，这个域就是该曲线的全集，但简单例子表明，其逆理不成立。还可指出，在欧氏平面上，可以这样界定闭曲线：它构成有界连续统，把平面划分成以它为共同边界的至少两个域。这个定义和推广了的旧定义不同；闭曲线的旧定义（舍恩弗列斯的）是：在一般闭曲线中，把平面划为刚好两个域的，称为“正则的”。（参看上面亚历山大罗夫的文章）。］

穷序列的闭多边形 P_1'，P_2'，P_3'，…其中每一个把后一个包围在内，但在它们外部，都没有点集的点(图 27.16)。若当假定已经知道平面上每个闭多边形总把平面划分为两个域，但这后来才有证明[多边形定理]①。

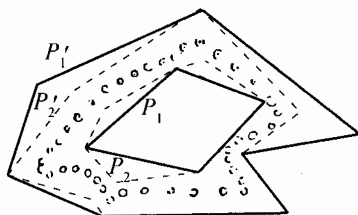


图 27.16

在这个思路下，再证明，每个不属于点集的点都或者在诸多边形 P 中某一个的内部或者在诸多边形 P' 中某一个的外部②，最后，所给点集本身就成了两个多边形序列 P_ν 和 P_ν' 的聚点。

对经验曲线的直观构成这个证明的指路牌。但每个具体论断逻辑上都扎根在公理中，而公理的核心则是现代实数概念。其结果可以归纳为：适用于诸多边形(周边)的东西也适用于诸多边形(周边)的极限图像。

可以把若当定理和前面指出的连续函数的一项性质相提并论：即连续函数要经过已指定的两个函数值之间的一切值。与此类似，闭若当曲线构成它的内部和外部的无间隙的界线。

由上面的讨论可以看出，在连通和分割方面，我们的若当点集是可以获得曲线这个称号的，但它是否具有和普通曲线的其他性质相类似的性质呢③？换句话说，对若当曲线是否可以谈论弧长、切线、

① [参看 B. V. Kerékjártó, *Vorlesungen über Topologie*(《拓扑学讲义》)，第 1 卷，柏林，1923 年。关于若当曲线定理的较新证明，可参看 G. Feigl, *Über einige Eigenschaften der einfachen stetigen Kurven*(法格尔：《关于简单连续曲线的性质》)第一部分。*Math. Zschr.* 第 27 卷，1927 年，第 161 页。]

② 原文是每个不属于点集的点都在诸多边形 P 或 P' 中一个的内部。——中译者

③ [上面没有提到近代拓扑学对处理曲线概念所起的卓越作用，诸如"可达性"，"无阻隔性"④，"局部连通"，"不可约连续统"等等。关于它们的研究，请参阅罗森塔尔在百科全书中的报告以及 Kerékjártó 的书。]

④ 上面底注中的"无阻隔性"是 Unbewalltheit 的暂译。设 P 为平面域 G 边界上一点。若 P 的任意邻域 U 里，总有 P 的邻域 V，使得在 V 内的每两个边界点都可以用一条既在 G 内又在 U 内的若当曲线相连，则 G 称为在 P 无阻隔。——中译者

曲率半径等等，或者对 φ, ψ 还要加上什么限制，才能谈论这些？（当然，我们必须对精确数学的点集提出相应的准确定义，而不是像在经验领域那样只提出近似的要求。）

我们先考虑关于若当曲线的弧长问题。

我们把 t 的有穷节 $a \leqslant t \leqslant b$ 任意地分成子节 $\Delta_1 t$, $\Delta_2 t$, …，并得到 x, y 的相应的增值 $\Delta_1 x$, $\Delta_1 y$; $\Delta_2 x$, $\Delta_2 y$; …把若当点集上的对应点用直线段按次序连起来，则这些线段的长是

$$\sqrt{\Delta_1 x^2 + \Delta_1 y^2}, \sqrt{\Delta_2 x^2 + \Delta_2 y^2}, \cdots$$

现在，设我们以任何方式把节 $a \leqslant t \leqslant b$ 分成子节的数目越来越大，同时各子节越来越小，而这样作出的多边形的长

$$\sum_{\nu} \sqrt{\Delta_\nu x^2 + \Delta_\nu y^2}$$

总有唯一的极限值，这时我们就把这个极限值作为点集的弧长。

从皮亚诺曲线的例子可以看出，这样的极限不总是存在的。对于该曲线，在引进参数 t 后，曲线弧长就不受其多重点影响，但 C_1 的长是 $\sqrt{2}$，C_2 的长是 $3\sqrt{2}$，C_3 的长是 $9\sqrt{2}$，…，C_n 的是长 $3^{n-1}\sqrt{2}$，…，故接连我们点集的关节点的多边形 C_n，随着边数增长不会有有穷的极限，它将超越一切极限。

在若当点集的款，也必须对函数 φ 和 ψ 加上一些限制，才能得到弧长。我在这里不给出证明，只指出其关键所在。

为此，我必须引进现代函数论用得很多的一个概念，即有界变差函数。已给在节 $a \leqslant x \leqslant b$ 上的函数 $f(x)$，若对于节的任意细分 $a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$，总和

$$\begin{aligned} & | f(x_1) - f(x_0) | + | f(x_2) - f(x_1) | + \cdots \\ & \quad + | f(x_n) - f(x_{n-1}) | \end{aligned}$$

总保持在一定的有穷值 A 之下，则在该节内，$f(x)$ 就称为有界变差

函数。

我们还要问,有没有我们知道的非有界变差函数?请回想讨论皮亚诺曲线所遇到的函数 φ 和 ψ,它们都是直锯齿形曲线的极限。几何上显示:对它们,当 n 增大时,$\sum_{1}^{n}|f(x_\nu)-f(x_{\nu-1})|$ 要超越一切界限①。

关于若当曲线的弧长,C·若当证明了,按照我们的定义,若当曲线有弧长的充要条件是 φ 和 ψ 都是有界变差函数。

§27.7 更狭义的曲线概念:正则曲线

若进一步问,什么时候若当曲线有切线或密切圆,则我们至少可以给出一个充分条件。我们的曲线在一个已给节的每点有切线和曲率半径的充分条件是,φ 和 ψ 在该节内可微两次②。

若我们的点集满足所有如下已举出的条件:

(1) φ 和 ψ 在节里连续;

(2) 没有多重点;

(3) φ 和 ψ 是有界变差函数;

(4) 有尽多的 ν 次 $(\nu\geq 2)$ 可微。

则对照经验曲线的类似性质,该点集就直截了当地称为曲线,若要强调它和诸如皮亚诺曲线的区别,我们就称之为正则曲线。或者更准确地,我们把这样界定的点集首先称为正则曲线段,这样,一条正则曲线就由有尽多个正则曲线段连接而成,因而可能有(有尽多个)

---

① [现在,经过N·维纳对理论物理的研究(白光研究),非有界变差函数有着重要作用。参看N·维纳:*Verallgemeinerte trigonometrische Entwicklungen*(《推广了的三角展开式》)。*Nachr. der Ges. der Wiss. zu Göttingen aus dem Jahre*, 1925(柏林,1926年,第151—158页)。]

② [这不是必要条件。因为在 φ 和 ψ 对参数 t 两次可微的情况下,我们还可以把 t 代以另一个函数 $t_1=f(t)$,使得 φ 和 ψ 对 t_1 并非两次可微。我们只需用一个单调连续而不可微的函数作为 $f(t)$,结果就成这样。]

重点。

我还补充一点：有时候我们把以前称为“光滑”的曲线也叫做正则曲线，那样的曲线在它每点有切线，而且当一点沿曲线连续移动时，切线也连续转动（图27.17），这比曲率半径的要求要少些。我们宁愿给正则曲线的定义留点弹性，使得对 φ 和 ψ 的一阶和高阶导数的要求可以适应不同的具体情况而有差别。此外，我还要回顾一下，在这个讲演的第一篇里，和这里讨论的不同的只是，在那里，我们限于用 $y = f(x)$ 表示的点集。这里采用正则曲线这样的术语，在那里则用“合理”函数。

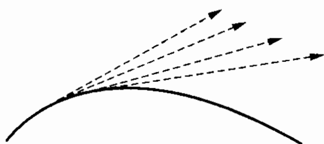


图 27.17

§27.8 用正则理想曲线近似表示直观曲线

我们还要问：这里在精确数学中所界定的图像和经验领域中直接称为“曲线”的图像之间，关系如何？由于在两个领域里都用了曲线这个词，为了避免用词上的不确切性，也为了我可以说得简短些，以下在精确数学中，我们将用理想曲线这个词。有时也用“正则曲线”；在经验领域中，则用经验曲线。

我指出：对于每条经验曲线，总可以设想有一条理想曲线（用 $y = f(x)$ 的形式表示）具有前者的主要性质。这里的意思是，在经验许可的准确度范围内，尽可能地有相同的性质。

我采用以下方法来证实我的论断。设想所给经验曲线可以分成有尽多段，每段可以在某个适当的坐标系下用公式 $y = f(x)$ 充分近似地表示——包括方向，有时还有曲率——（参看第 60—64 页的论述）。于是有有尽多个坐标系

$$x_1y_1,\ x_2y_2,\ \cdots,\ x_ny_n,$$

在这些坐标系里面分别有理想曲线中正则的一段

$$y_1=f_1(x_1),\ y_2=f_2(x_2),\cdots,\ y_n=f_n(x_n);$$

这些曲线段(通过坐标系的相互关系)首尾相接以构成一个连通的整体。现在,我们引进一个单一的坐标系 XY 和一个参数 t,这个参数的值就分布在有关节的 n 个子节里。当 t 在第一个子节里时,就把它用 t_1 表示,在第二个子节里时,用 t_2 表示,等等。于是 n 个公式

$$y_1=f_1(x_1),\ y_2=f_2(x_2),\ \cdots,\ y_n=f_n(x_n)$$

就可以合为单一的一对公式

$$x=\varphi(t),\ y=\psi(t);$$

当 $t=t_1$, $t=t_2$, …时,这对公式依次代表 $y_1=f_1(x_1)$, $y_2=f_2(x_2)$,… 同时,在这些子节里,φ, ψ对 t 可微次数和 f_1 对 x_1, f_2 对 x_2,… 可微次数相同。为此,最简单的方法是令

$$t_1=\alpha_1x_1+\beta_1,\ t_2=\alpha_2x_2+\beta_2,\cdots。$$

再确定这里的常数 α, β,使 t_1 的终点值(对应于 x_1 的右端点或左端点值)等于 t_2 的始点值(对应于 x_2 的左端点或右端点值),同样,也使 t_2 的终点值等于 t_3 的始点值等等。

这样扼要说明的步骤,其结果还存在一种可能性,即所得到的不同的理想曲线段 $y_1=f_1(x_1)$, $y_2=f_2(x_2)$,…在它们相接处会有不同方向或曲率。与此相应,所给经验曲线就会有有尽多处,在那些地方,它的经验的方向或曲率有跳跃式的变动。所得理想曲线 $x=\varphi(t)$, $y=\psi(t)$ 自然也就有相应的奇异性。

§ 27.9　理想曲线的可感知性

我们还要对上面所讨论的问题的逆问题作些补充;我们要问:

在经验领域中要考虑的是正则的理想曲线 $x=\varphi(t)$，$y=\psi(t)$ 的哪些性质?

显然，要考虑的是：当 φ 和 ψ 在一定范围(这范围决定于具体情况)内任意变动时保持不变的那些性质，或者说，它们是近似数学所讨论的性质。这种说法可以直观地表述如下：设想在曲线 $x=\varphi(t)$，$y=\psi(t)$ 上每一点作一个具有已经给定的半径为 ρ 的小圆。这样所作的圆盘的集体覆盖平面上某一个带，而在经验领域里，要考虑的是这样确定的带所呈现的性质而不是理想曲线自身的性质。

在上面已获得认识的基础上，现在又回到我们讲演开始所涉及的问题：即我们对空间感觉的准确性问题。

假定在一个坐标系 x，y 下，一方面已给定一个圆；另一方面已经通过公式 $x=\varphi(t)$，$y=\psi(t)$ 给定一条皮亚诺曲线。凭我们的空间感觉，我们有能力从原则上区别这两款吗? 我认为不能。在这两款中的每一款，我们都可以设想，集中注意力于一个具体 t 值在图像上所确定的一个具体位置，但总是在有限度的准确度之内。我们还可以类似地设想，有确定的邻近点，并在有限度的准确度内感觉到这两点连线的方向；可是，无论在这一款或那一款，都不能使我们的感觉能力达到理想图像本身，我们驾驭理想图形的数学思考能力来自空间感觉并受它引导，但这种思考最终还要在公理的基础上，以公式 $x=\varphi(t)$，$y=\psi(t)$ 所确定的规律为依据(而公理也是来自直接感觉能力)。

这就是 1873 年我的论点[①]。我希望生理学和心理学能澄清这个问题，只是有一个前提：必须通晓精确数学的现代发展(例如像本讲演所阐述的那样)并且吃透了它，如果不是深入考虑了像魏尔斯特拉斯曲线和皮亚诺曲线那样不可微的理想曲线的具体例子，就不可能对这类理想曲线和普通可微理想曲线的区别进行哲学上的讨论。

---

① 参看前言第 ii 页所引文献。

§ 27.10 特殊理想曲线：解析曲线与代数曲线，代数曲线的格拉斯曼几何产生法

现在转而略微谈谈精确数学里优先讨论的主要类型的理想曲线。

(1) 什么时候我们把一条正则理想曲线叫做解析曲线，即在正则理想曲线的定义中，要加上什么样的限制条件，才能把它称为解析曲线？一条正则解析理想曲线将简称为解析曲线。

我们要说：当 x 和 y 可以用 t 的收敛幂级数表示时，正则理想曲线就称为解析曲线①。

可是，我们既然早就知道，解析曲线远远不能包括一切曲线，为什么数学家总还是集中力量于解析曲线的研究②？

原因在于解析曲线所具有的那种美妙的性质：首先是它可以推广到复领域 $t=u+iv$，这样就涉及整个(复变)函数论(其幂级数在 uv 平面上一个区域里收敛)。特别是存在着“解析开拓”到新的定义域的可能性，因而只要该区域含有实轴的一段，原来一段曲线就可能延伸以得到新的一段曲线。总之，可以说，一条解析曲线连同其解析开拓，构成一个合乎规律的整体；它用某个幂级数 $x(t)$，$y(t)$表示，而这两个级数只需在一个任意小的区域内收敛，曲线整体就完全确定了。这里所说的规律性——最小一段决定整体——就是数学家们如此喜爱解析曲线的根由。为了说得确切些，我还要补充如下：

一条解析曲线还可以具有各种不同的奇点；若假定在 φ 和 ψ 的级数展开

$$x=\varphi(t)=a_0+a_1t+a_2t^2+\cdots$$
$$y=\psi(t)=b_0+b_1t+b_2t^2+\cdots$$

① [已经知道，这个定义和第 76 页关于解析函数的定义等价。]

② [现在，这个问题已不像讲演发表时那样突出。]

中，所有从零到 $n-1$ 次($n\geqslant 2$) 项都不出现，然后我们考虑 $t=0$ 处的情况，就得到这种奇点。

(2) 现在我们把曲线概念进一步限制，从解析曲线进入代数曲线。

用文字表达：若解析函数 $x=\varphi(t)$，$y=\psi(t)$ 恒等地满足一个代数恒等方程 $F(\varphi,\psi)=0$，则解析曲线称为代数曲线。

逆定理也是成立的：已给一个代数方程 $F(x, y)=0$，则在每点 x_0，y_0 ① 的邻近可以引进一个辅助变数 t，使 $x=\varphi(t)$，$y=\psi(t)$，其中 φ 和 ψ 都是 t 的收敛幂级数(单值化)。对此，我不作证明，只是作些引申而不涉及其细节；因为我在这里只是要对事物作一般的概括。

现在进一步的问题是：对代数曲线，除了这个形式上的定义之外，能否给出一个更具体的定义，即直观上可以掌握的定义？实际上，这从函数论的角度和几何角度都是办得到的。

首先是从函数论角度。假定消去 t 以得到函数 $y=f(x)$。问题是：对于具体的 x 值，其幂级数是否有什么特性，可以把代数函数区别于其他解析函数。在这里，x 自然要作为复变数，因为那样才能通过函数论的语言来按其自然法则求得其表达式。

实际上，从函数论的观点，在 y 作为 x 的代数函数的定义中，我们要求：

(a) 对于具体的 x 值，只有有尽多的 p 个 y 值；

(b) 在每处 x_0 的邻近，y 可以展成 $(x-x_0)^{\frac{n}{p}}$ 的幂级数，其中负幂只能有有尽多个(这样就排除了所谓的本性奇点)。具体的论证当然属于函数论的课程。

比函数论角度更有趣味又更易懂的是和几何相联系的说明方式。关于代数曲线的几何特征，早在 1850 年，H·格拉斯曼就给出

① 假定 $F(x_0, y_0)=0$。——中译者

了决定性的答案[①]，可惜他的论述至今未在教材内出现。我下面采用名词"直尺机械"来叙述格拉斯曼定理：

若一条曲线可以通过直尺机械产生，它就是代数曲线。

首先，什么是直尺机械？

一个直尺机械是由部分固定、部分作运动的直线和点构成的，其中作运动的直线必须经过其中某些(不一定是固定的)点，而作运动的点必须在其中某些(不一定是固定的)直线上。

为了具体地阐明这个问题，设想已给定一定数目的直线和点。格拉斯曼直尺作图法的思路是，从已给固定的和作运动的直线和点，只利用直尺以得到新的点和直线。若规定作图时要由最初的元素出发，则直尺作图的机制就构成直尺机械。

作为例子，我现在说明如何通过一种直尺机制来作圆锥截线[②]。

设在平面上给定 3 个点 a，b，c 和两条直线 B，Γ(图 27.18)。若规定平面上一个动点 p 受到下述条件制约，则点 p 描绘一条圆锥曲线：直线 $\overline{bp}$ 和 $\overline{cp}$ 依次和直线 B，Γ 交在 β 和 γ，a 同 β，γ 共线(麦克劳林产生法)。

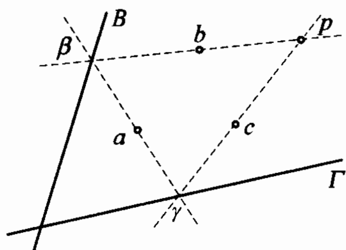


图 27.18

设想用刚体的线，最好是用缝衣针，制成一项装置，使它按照所述的方式运转，则点 p 就沿一条圆锥截线运动。这就是一个直尺机械。一般地，每一个这样的机械都有如下的性质：由点 p 通过直尺作图得到的 3 点在一条直线上，或者由点 p 通过直尺作图得到 3 条直线经过一点。不过，直尺机械只有理论意义，因为制作时不能准确地实现对它的要求，即使能制成，它还总是晃动的。

① [参看 *H. Grassmanns Gesammelte Werke*(《格拉斯曼论文集》第 1 卷 1，第 245—248页，莱比锡，1894 年，和第 2 卷 1，论文 2—7，14，18，1904 年)。]

② 即二次曲线。——中译者

格拉斯曼利用这个思路来给出代数曲线的几何定义。

在圆锥截线，即二次代数曲线的产生中，点 p 用了两次。格拉斯曼发现，若把机械组织好，使 p 点用到 n 次，就得到 n 次代数曲线①。

他并不就此停下来，他还证明了逆理：每一条 n 次代数曲线 $F(x, y) = 0$ 都可以通过（也许很复杂的）直尺机械得到，其中作图时，点 (x, y) 将用到 n 次。他证明时，把方程 $F(x, y) = 0$ 的要求直接转化成那个机械的构造规范。

有了这个定理之后，就立刻有了上面提到过的关于代数曲线的格拉斯曼定义：一条 n 次代数曲线是可以通过一个 n 次直尺机械产生的曲线。

我们再一次问，代数曲线使人感兴趣的是什么？

当然（像一般解析曲线那样）它的整体被任意小的一段所确定。此外，它还有其他关于其整体的简单性质。我们只举出其中最重要的：

它有有尽多个分支，而且对于它，贝佐定理成立：即一条 n 次代数曲线与一条直线交于 n 个点，而且两条不同的代数曲线 C_m 和 C_n 相交于 mn 个点（假定我们采取了正确的计数方法：即包括虚交点；有时交点会在无穷远处；对于多重点，按其相重数计算；还要考虑到两条曲线可能有共同的分支）。

在这一般性定理的基础上，特别是对于从一次到四次的曲线，导致了非常美丽的结果，使人们对它们的探讨得到富有魅力的乐趣和审美上的满足。

因此，对代数曲线的研究有些时候曾成为数学学习中的爱好对象。但若像许多教材那样，光讨论代数曲线，则只应当认为，那是片

---

① [n 次代数曲线可以通过直尺机械产生这个特性不能理解为，似乎每条 n 次代数曲线都必须是这样的曲线，它们单纯借助于直尺就能产生。按照直尺机械的关联条件，只通过直尺来确定一个点 p 一般来说是不可能的。]

面的。人们会由此获得一种印象，似乎只有这样的曲线了；可是正弦曲线和螺旋线出现得远为频繁，而且至少同样地美妙。

比代数曲线更为特殊的是有理曲线。这时人们仍然由参数方程 $x=\varphi(t)$，$y=\psi(t)$ 着手，但这里的 φ 和 ψ 是 t 的有理函数。对有理函数的研究，人们探讨得也不少。

§ 27.11 用理想图形表现经验图形；佩雷观点

对精确数学领域里曲线中的特殊类型进行阐述之后，我们认真地提出那个一般性的问题：

这种种理想曲线对于应用起什么作用?

一种老成见认为，在应用中遇到的都是解析曲线。我们已反复强调与此相反的观念：在应用中所遇到的根本不是理想曲线而是和它们近似的东西(就是带)。但是，为了描述一条经验曲线或者理想化的现象，人们也用到理想曲线，而且总是用正则曲线(即 φ 和 ψ 常常是有尽多次可微，而且其中最后一个导数也许还是逐段单调的)。如果相反地认为必须采用解析曲线，那是形而上学。我把那种不从直接经验出发的作法称为形而上学，因为人们一厢情愿地只使用解析曲线。

不过，尽管我们认为在应用中，解析曲线、代数曲线、有理曲线等都不会出现，但好在情况却是，最简单的正则曲线的确是解析的。

例如，当我们已经得到

$$x=a_0+a_1t+a_2t^2,$$

$$y=b_0+b_1t+b_2t^2,$$

时(在经验领域中，为了表现一段具体曲线，人们往往简单地使用这样的公式)，我们就遇到一条有理曲线。在应用领域中，我们也不能

避免用到 $y=\sin x$, $y=e^x$, …尽管我们需要作出修正,而修正后曲线的解析性已变了样。因此,我们可以对上面的论点作如下的补充:

在经验领域里,解析曲线、代数曲线、有理曲线只是在这样的意义下有用:它们是人们所了解的最简单的正则曲线,而在近似地表现经验所取得数据时,也会经常地使用它们。

在这里,我想对 J · 佩雷(Perry)的书,*The calculus for engineers*(《工程师用的微积分》)[①]说几句话。在这本书中,我上面所谈的观点得到有特色的发挥。

这本书的特点是,作者为应用科学工作者写书,因而完全站在他们的立场上。他不谈一般的曲线概念等等,而开始就只讨论应用中的 3 种基本函数($y=x^n$, $y=e^x$, $y=\sin x$)。他指出,只要运用由这 3 个函数经过有尽多次四则运算所获得的函数,例如

$$y=\frac{a+bx+cx^2}{d+ex+fx^2},$$

等等,就可以足够准确地解决一切普通的应用课题。

佩雷的这种做法被反复指责为不科学的。但在其所涉及的领域里,根本不是如此,我现在所进行的讲演,其目的正是要采取包括理想领域和经验领域在内的观点。而在这个构思中包含了佩雷的观点。

---

① 伦敦,1897 年,被弗里克和聚希廷(Fr. Süchting)译成德文[②],莱比锡,1902 年,第四版,1923 年——[关于佩雷运动,可参考本讲演第二卷第 232—237,280 页*]

② 在中译本里,这部分删掉了。——中译者

第二十八章 继续从精确理论观点讨论平面几何

§28.1 对两个相切圆的相继反演

现在我们捡起前面曾经提到过的一种思路(因而我再次转入完全理想的领域)。我们已把整个论证和解析几何的思想结构相联系，因而把一切放在现代实数概念的最终基础之上。如果要想把至今为止所谈论的内容从这种论述方式中解脱出来，把它放在纯几何的基础上，就必须把相当于引进现代实数概念的一个公理放在首位；这个公理可以叙述为：若有一个无尽序列的闭节(线段、曲线段、平面区域、空间部分)，其中每个含在前一个之内，而且它们无限制地缩小，就有唯一的一点属于所有的节。这个点就被那个节序列唯一地确定。这个以前曾利用过的公理，叫做节套公理①。

这个公理将选取为下面所论内容的理论基础，我们以前已看到，在自守函数理论中，曾利用它来界定与讨论一些具有集合论中所涉及的令人瞩目性质的点集。

现在我要对导致非解析曲线的几何产生原则略加修改。在这里，产生过程是纯几何性的，这样，就同时避免给人一种印象，似乎讨

---

① [容易看出，为什么在上面的公理中要求各节是闭的。若 n 是自然数序列，则对于无限制地缩小的开节序列 $0<x<\frac{1}{n}$，没有任何点会属于所有的节。0 点不属于序列中任何的节。]

论非解析曲线是人为地强加于几何中的。

前面我们曾经从平面上 3 个或更多个圆出发，其中每个都在其他几个之外，并把以它们为边界的区，通过对它们的反演，以不断得到和该区等价的新区（图 28.1）。那样，我们得到等价区构成的网，这个网覆盖除一个无穷集合的极限点以外的整个平面，然后提出这样的问题：对于这个极限点集，我们能说些什么？我们发现，它是无处稠密的，却又是完备的，并有着连续统的势。

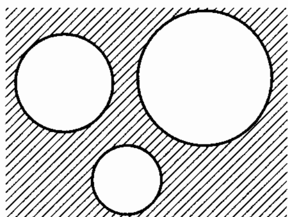


图 28.1

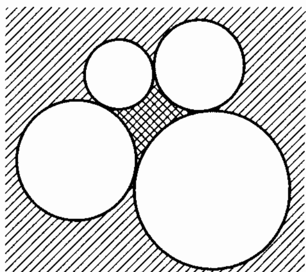


图 28.2

在这里，我们改变一下，假定这些圆构成循环相切的序列。但和以前一样的是，每两个圆都不交叉。这样，初始区就分成两个以圆弧为边的 n 边形，一个是内区，一个是外区，而两个 n 边形的角全都是零。可以参考图 28.2，其中一个区延伸到无穷远处，但这并不重要，因为经过一次反演，它就变成一个有界区。现在，我们让两个具有零角的圆弧多边形对其边界圆相继地作无尽多次反演，并考察如此得到的极限点集。我们将发现，由极限点及其聚点构成的集合，是一条闭若当曲线，它一般地不是解析的，这里所提到的附加词"一般地"，将在讨论中加以说明。这条若当曲线就把从初始的外多边形及其像所构成的网和内多边形及其像所构成的网分隔开来。

在这里，我也要再次强调，对若干个循环相切的初始圆把初始区及其等价区不断作反演，这种构想来源于物理。当人们在静电学里讨论若干个圆柱上的电荷的平衡分布时，就恰好要考虑上述那样的

无尽次的反演，因此，只要人们的实践探究进行得充分深入，就不能避免非解析曲线。

先设有两个互相外切的圆 K_1 和 K_2，这时当然谈不上循环相切，我们只有一个初始区，它是一个圆弧二边形，有两个零角，在图 28.3 里，这个区也像以前那样，用 1 表示。若把 1 对 K_1 反演，就得到以 K_1 为部分边的一个镰刀状的区，这样所得到的区和对 K_1 的反演都用 S_1 表示；同样，对 K_2 反演，就得到 K_2 内一个镰刀状区 S_2（图 28.4）。在这里，我们以前所说的进展原则仍然适用，即除了开始的反演外，我们把已得到的每个区继续对它的内边界圆作反演，以前证明过，这个原则已经包括了对初始区作反演 S_1 和 S_2 的一切组合在内。我们对这个原则还给出了一个推广了的形式，即对于每一个新得到的边界圆，不但把以它为边界的区作反演，而且把已经得出的一切区作反演。在我们所讨论的款，例如已经得到的 3 个区 1，S_1，S_2（图 28.4）后，立刻可以对 K_2 作反演以得到 3 个新区，等等。现在，无论按这个或那个方式持续进行下去，就得到定理：对于我们的等价区有一个唯一的极限点，它和两个初始圆的切点重合。

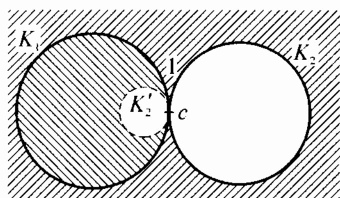


图 28.3

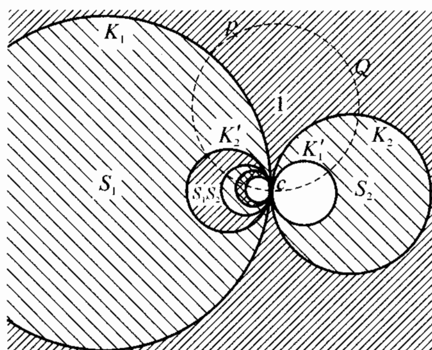


图 28.4

一个有意义的问题是，已给域 1 里任意点 p，考察一下，p 的无尽多个等价点是怎样趋近极限点的？取经过 p，而同两个初始圆正

交的圆 Q，我们知道，对于我们的一切映射 S_1，S_2，…，Q 变为自己。因为它经过圆，K_1，K_2 的切点 c；c 以及 Q 在 c 的切线对于一切 S_1，S_2，…都不变；此外，Q 和 K_1 的交点对于 S_1，Q 和 K_2 的交点对于 S_2 都不变，结果是，经过反演，由 p 得到的一切等价点都在正交圆 Q 上。我们可以把 Q 看成是 p 跳跃地趋于极限点 c 的"轨线"。

特别有趣的是，把极限点和平面上任意点用直线相连时，因为当对 p 运用我们的进展原则中的反演时，p 在 Q 上趋于 c，p 和 c 的连线必然要趋于一个极限位置，即两个圆的"中心线"①(图 28.5)。

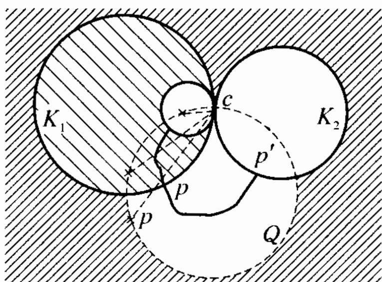


图 28.5

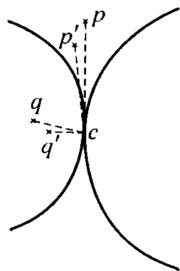


图 28.6

现在考虑按下述方法做成的闭曲线：在初始区 1 里，由 K_1 到 K_2 作任意一条不含点 c 的曲线段(可以是非解析的)②，然后接上由它通过我们的反演所得到的一切曲线段③，最后加上点 c。我们将证明，这样得到的闭曲线在点 c 有确定的切线，即中心线。

这个定理是上面所述引理的一个简单结论。为了证实这一点，在曲线上(图 28.6)选取趋于 c 的任意点序列 q，q'，…。一切点 q，

① 即连接相切圆 K_1，K_2 的中心的直线，它也是 Q 在点 c 的切线。——中译者

② [这当然是指用方程 $x=\varphi(t)$，$y=\psi(t)$ 确定的一条曲线段；至于按照所说方法，的确出现一条闭曲线(一个圆的拓扑像)是不难证明的]。

③ 根据前面所说，只需把初始区里所取曲线段分别对 K_1 和 K_2 反演，再取这两个反演像分别对 S_1 和 S_2 的内边界圆作反演，这样不断继续下去，就得到那段曲线段的一切等价曲线段。——中译者

q'，…(它们逐步趋近于 c)在初始区中，都有一个等价点。我们可以把 q，q'，…挤进 p(和 p'，参看图 28.5)的等价点 p_v(和 p_v')之间①。因此直线 cq，cq'，…的极限位置和 cp_v(和 cp_v')的极限位置相同，于是根据上述引理，cq，cq'，…的极限位置就是中心线。

§28.2　对3个循环相切圆的相继反演("模图形")

现在由两个初始圆的简单款进入 3 个圆的款。为方便起见，我们选取 3 个大小相同的圆，于是初始区就由两个具零角的圆弧三边形所构成(图 28.7)。我们知道，这样的 3 个圆总有一个唯一的共同正交圆 Q。它经过 3 个切点。对于所有将要遇到的反演，它都不变，因而可以用来帮助探讨我们的图像。

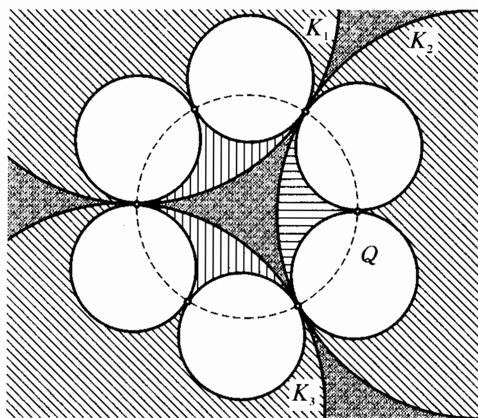


图 28.7

为了得到两个初始区的一切等价区，我们只需作出在正交圆 Q 内部的等价区，然后把这样得到的全部等价区对 Q 作反演。可以这

① 图 28.5 中用 p 和 p' 依次表示这条曲线段在 K_1 和 K_2 上的端点。——中译者

样做的根据是早先指出过的定理：若对于另一个圆 Q，圆 K 的反演像是 K'，则对 K 互为反演像的两个图形，经过对 Q 的反演，变成对于 K' 互为反演像的图形。因此，可以把我们的作图限于在 Q 的内部①。由于有 3 个旁三边形和初始区相接，这 4 个三边形在一起构成一个圆弧六边形，它的尖角和 Q 正交。根据我们的进展原则，可以把 3 个旁三边形分别对它们的内边作反演，或者把整个六边形对它的 6 个边分别作反演以得到一个三十边形。你们可以看到，继续下去，结果将会怎样：由初始区所得到的区把 Q 的内部覆盖得越来越满，所得区的尖角越来越多，并且都和 Q 正交。当这个步骤越发进行下去时，初始区的等价区越来越密切接近 Q，而 Q 的每点成为尖角的聚点。对此，我们将作进一步的阐明。

这里所描述的图形，对于研究椭圆模函数的人们是熟知的；在一定意义上，这是几何从函数论接受的礼物。因为尽管这个无限继续下去的步骤完全可以从纯几何观点出发得到，但从历史发展过程来看，是由于椭圆模函数理论的需求，人们才开始考察这样的图形的。不过，这图形先已出现在 v·施陶特(Staut)的工作中(当他对于二次曲线上 3 个点，不断求其第四调和点时)，但 v·施陶特没有在直观上加以揭示。

我对 3 个相切圆的图作较深入的考察，以便突出其中的一些细节，使 4 个相切圆的图形的研究变得容易些。

为了得到对称图形，我选取半径相等的相切圆，使它们的切点成为正交圆 Q 的内接等边三边形的尖角。我们或者先把由内区通过持续反演所得的部分图形对 Q 作反演以得到整个图形，或者先把由外区通过持续反演所得部分图形对 Q 作反演。

现在我们考察，这样所得的圆弧三角形的切点在正交圆上是如

① 对 3 个初始圆中的每一个，作其他两个的反演像，就得到 6 个小圆，它们都和 Q 正交；而且每个初始圆内都有这 6 个中的两个，它们和初始圆相切，也彼此相切(图 28.7)。——中译者

何积聚起来的。我们本来有 3 个切点,然后有 6 个,12 个,等等,所以可以说:当作图步骤继续时,整个正交圆越来越密地被新出现的圆的切点所布满。

这里就提出下一个问题:对于切点所构成的集能有什么结论?容易看出,在切点集合中,每个切点是其他切点的聚点,即这个点集在正交圆上是自稠密的。

为了证明这个定理,考察任意两个在我们的步骤中所得的相切圆"1"和圆"2";令它们在 Q 上的切点为 c。设想按我们通常方法,把这个图形(图 28.8)不断地作反演,首先对圆"1"和圆"2"作反演,然后把整个所得图形对新的边界圆作反演,等等。这样,在第一次反演之前本来是两个圆的切点就成为无尽多个圆的公切点。随着反演步骤的进行,这些无尽多个圆也无限制地变小,而且它们每一个和正交圆又有另一个交点,而且还是一个新的切点。图 28.8 表明这些新切点是如何属于相连接的圆弧三角形的。图 28.9 说明正交圆退化为直线时的特款。

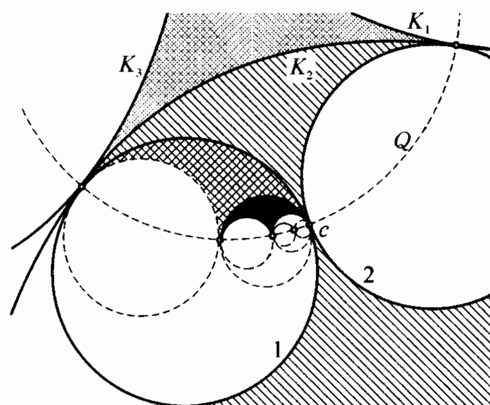


图 28.8

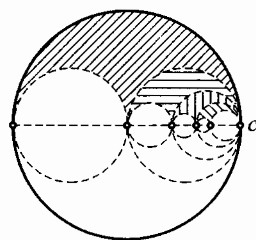


图 28.9

由此可知点 c 的每一个无论多小的圆形邻域里总有无穷多个切点;因此 c 是聚点,于是切点的集合是自稠密的。

作为这个点集的另一个性质,我要指出,在正交圆圆周上,点集

是处处稠密的，即在圆周上任意小的一部分总含有切点。证明如下：我们的初始图形是 3 个圆构成的闭链(图 28.7)，它们沿着正交圆排定。把这个链对这些圆作反演，就得到 6 个圆，它们也沿正交圆排列成一个闭链。一般地，对每一个这样的链中所有的圆作反演，就得到一个新链，其中的圆越来越多，也越来越小。现在我们在正交圆的圆周上取一段弧，它的弦长是 δ，我们把反演步骤继续下去，直到所得的圆链中的圆的直径都小于 δ。这时，无论 δ 多么小，所取圆弧上肯定有切点，因而当反演步骤无限制地进行下去时，这个圆弧上就有无尽多个切点。这同时也表明，正交圆上的一切非切点——我们将说明这样非切点存在——也都是切点的聚点。

在正交圆 Q 上，处处稠密的切点和 Q 的其余点之间的关系完全像数轴上有理点和无理点的关系。在数轴上"有理点"也是处处稠密的，而且是可数的，而"无理"点则显示为有理点的聚点。如果我们想到，按戴德金的做法，一个无理数可以通过具有某种性质的对有理数的分割来界定，则目前的对比就更清楚了：正交圆上每一个非切点确定切点集合的一个分割，因而也被这个分割所界定。

我的阐述不是很确切的；因为我只是联系着图形来谈，没有涉及数量关系。但容易把图形改变一下，使得它可以用算术来处理。

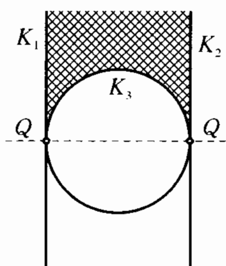


图 28.10

由于所要考察的是那些经过对整个图形的反演不变的性质，可以令初始区具有简单的形状，不会失去普遍性。我们把初始区对中心位于一个三边形某个尖角的圆作反演，这样就得到一个三边形，它的边界是两条平行直线和一个同它们相切的半圆(图 28.10)。正交圆变成一条经过两个有穷切点的直线 QQ，这就是通常在函数论里所给出的表现所涉及的关系的图(这样还便于运用坐标系)。

取直线 QQ 为 x 轴，选取原点和长度，使两个有穷尖角在 $x = 0$

和 $x=1$，于是整个图形就容易作出了。我们只要把最初图形在左右两方无限制地重复下去，再把这样得到的无尽多个三边形反演到所出现的半圆内(图 28.11)。这样，切点就都在 x 轴上，而且原先只是通过直接几何感觉所阐明的结论现在可以很容易地用算术来准确地计算。由所建立的公式可以证实，各切点的横标都是有理数，而且 x 轴上横标为有理数的点都是切点。因此，切点的集合和有理数集合不只是类似的，而且是同一的、特殊的，这表明切点构成可数集。

这是对三圆之款的补充。

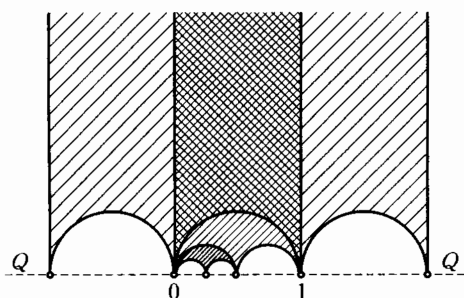


图 28.11

§ 28.3　4 个循环相切圆的标准款

我现在将讨论中的问题转入 4 个循环相切圆的款。在这里，从整体看和上面的款相类似。

先假定这 4 个圆有一个共同正交圆(图 28.12)。这样，作为初始区，可以先使用所出现的两个圆弧四边形中外面或里面的一个[①]。我们选取里面的一个，并进行反演。按我们的进展原则，先有一个由内四边形以及同它相连接的 4 个四边形所构成的 12 边形，然后有一个 36 边形等等。在这个完全对称的步骤中，正交圆上的切点数依次

① 外面的 1 个是含有无穷远点的那个。

是 4, 12, 36, 108, …我们就要问：在正交圆的圆周上，切点的分布如何？我们可以指出：

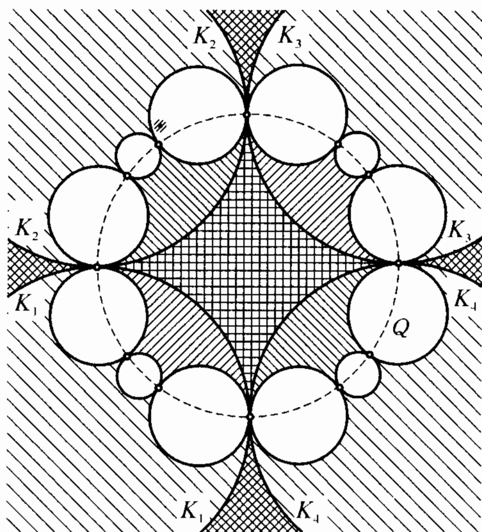


图 28.12

(1) 切点集合是可数的。

(2) 切点集合在正交圆上处处稠密。

现在，为了使我们的图形具有特殊的形式，借助于一个适当的反演，把一个切点变到无穷远处，所得到的是图 28.13，它可以用算术方法处理。

正交圆退化为一条直线，我们取它为 x 轴，4 个圆中有两个仍化为平行直线，其余两个则化为夹在它们之间而与 x 轴正交的圆。相继对那两条直线不断向左和向右作反射(即反演)，并对那两个圆作反演，就得到整个图形。仍令外边两个切点在 $x=0$ 和 $x=1$，并用 σ 表示第三个切点的横坐标 $(0<\sigma<1)$。在 x 轴上，所得切点集合不再和有理数重合，但这些切点的横坐标是 σ 的有理函数，考察这些有理函数的性质是有意义的。

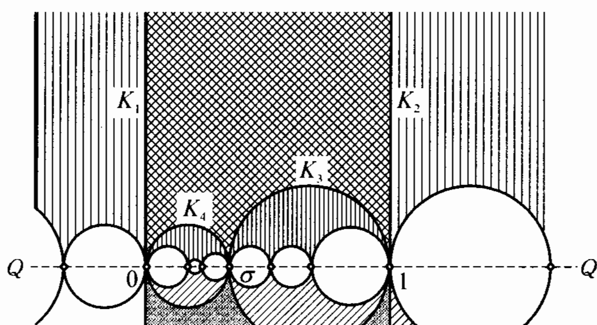


图 28.13

§28.4　4个循环相切圆的一般款

上面是4个初始圆有一个公共正交圆的特款。现在考虑一般款，它当然复杂得多。但这时利用以前的定理，仍然可以获得某种概略结论。设已给4个循环相切的圆。这时候，值得注意的是，那4个切点仍然在一个圆上，但这个圆一般不是正交圆(图28.14)。

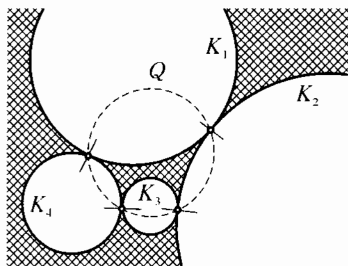


图 28.14

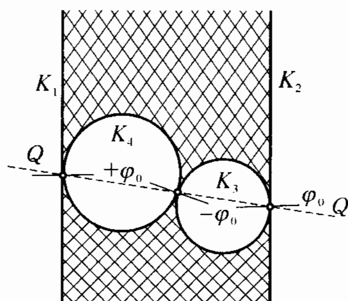


图 28.15

为了证明这个引理，我们运用已用过的技巧，通过对于以切点之一为圆心的一个圆的反演，把图形简化。这时，在这个反演中心相切的两个圆就化为两条平行直线 K_1 和 K_2，夹在它们中间的是另外两个圆 K_3 和 K_4，像图28.15所表示的那样。

为了证实原来图形中 4 个切点在一个圆上，就要证明图 28.15 中的 3 个切点在一条直线上。但根据初等几何，这是立刻看得出的，于是引理就证明了。这条直线叫做截线 Q。一般地，经过圆弧四边形 4 个顶点的圆就叫做截圆 Q。

与此同时，若在图 28.15 中的各切点作圆的法线（它们构成两对，分别交于 K_3 和 K_4 的中心），则我们的这些截线和法线或中心线（在这里，我们把法线也称为中心线，更能显示其特点）显然交于绝对值相等而异于 0 的角 $\pm\varphi_0$，其中的符号如图 28.15 所示。若 $\varphi_0=0$ 就得到以前所论的款。此外，容易看出，φ_0 总是小于 $\frac{\pi}{4}$。把这些结果搬回到原来的图形，我们就可以说：4 个切点在一个截圆 Q 上，它和切点处的中心线交于 $\pm\varphi_0$ 角，而由于 Q 不再是正交圆，φ_0 不等于 0，但总小于 $\frac{\pi}{4}$。利用这个定理，可以把原来的图形换成以后处理更为便利的形式。我们知道，若把一个复杂的图形代以一个有对称构造的，就容易了解得多。所以，下面我们画原始图形时，就从截圆 Q 画起，再选取它的一个内接长方形的顶点为切点（图 28.16）。

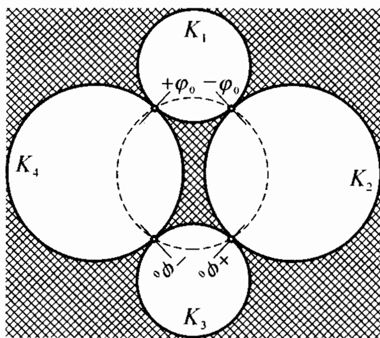


图 28.16

假定适当地选定异于 0 的 $\pm\varphi_0\left(\varphi_0<\frac{\pi}{4}\right)$ 角，按此画出中心线的方向，就容易倒转来作出 4 个“初始圆”。这样就得到两个圆弧四边形，一内一外；对于截圆 Q，它们并不互为反演像，但它们位置之间的关系却仍然像反演中那样。现在先把 Q 对 K_1 作反演以得到一个新的圆 Q'。由于 K_1 和 Q 的交点以及角 $+\varphi$ 和 $-\varphi$ 的绝对值不变，作出 Q' 变得容易了，但需注意要改变角的符号。其结果是，我们所着重考虑的，在

K_1 内部的一段弧 q'，相对于内圆弧四边形，其凹的一侧翻转了。同样，也容易作出 Q 对于 K_2，K_3，K_4 的反演像。这样，就得到圆 Q''，Q'''，Q^{IV}；设它们在 K_2，K_3，K_4 内部的弧为 q''，q'''，q^{IV}，像 q'那样，圆弧 q'''凹的一侧翻转了，而 q''和 q^{IV} 却是凸的一侧翻转了。于是，作了 4 次反演之后，截圆 Q 化为 4 个圆弧所构成的截曲线，这 4 个圆弧在它们的切点和中心线作 $\pm\varphi_0$ 角；容易看出，在它们相接处不是折点①（图 28.17）。

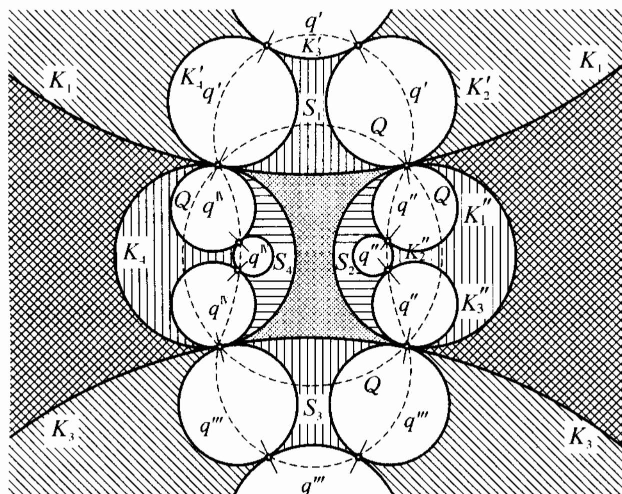


图 28.17

在画出这条新的截曲线之后，我们把 K_2，K_3，K_4 对 K_1 作反演，所得的新切点既在 q'上，又分别在圆 K_1 的中心和原来切点的连线上。在正确地画出新切点的中心线后，K_2，K_3，K_4 的反演像 K_2'，K_3'，K_4'就可以很快作出了。K_2'，K_3'，K_4'和 K_1 一起构成两个初始四边形的第一次反演像的边界，与此类似，可以作 K_1，K_3，K_4 对 K_2；K_1，K_2，K_4 对 K_3；以及 K_1，K_2，K_3 对 K_4 的反演。在利用截圆 Q 和圆弧 q'，q''，q'''，q^{IV} 作出以上容易理解的最简单的结果后，

① 即圆弧在相接处有相同的切线。——中译者

我们可以把所得图形用下面文字描述：

开始有一个截圆 Q 和它上面的 4 个切点以及过诸切点的中心线。由圆 Q 得到的 4 个圆弧构成一条截曲线，上面有 12 个切点，附有 12 条中心线，沿着截曲线上的 12 段有 12 个圆 K^{I}，K^{II}，…，K^{IV}，…这些把平面分割为一个内部，一个外部。

如何继续作图是明显的，取 12 个新圆中的一个，例如 K_2'。有两个圆弧四边形以 K_2' 的弧作为其 4 边之一。把它们对 K_2' 作反演，就得 3 个位于 K_2' 内的 3 个新圆①。它们之间的切点都位于 q' 对 K_2' 的反演像在 K_2' 内部的那一段 (q') 上（图 28.18）。上面用到的那两个四边形，除一边在 K_2' 上外，还有一边在 K_3' 上，一边在 K_4' 上。把那两个四边形分别对 K_3'，K_4' 反演，则在这两圆内部分别得到 3 个圆以及这 3 个圆的切点所在的一个圆弧。于是圆 Q' 的弧 q' 经过 3 个反演后被由 3 个圆弧构成的曲线所代替。这些圆弧相接处也不是折点，它和有关的中心线相错地作交角 $\pm\varphi_0$，现在，考虑在 K_2' 内部出现的两个圆弧四边形。把它们连同圆弧 (q') 对 K_2' 内的 3 个圆分别作反演，则在这 3 个

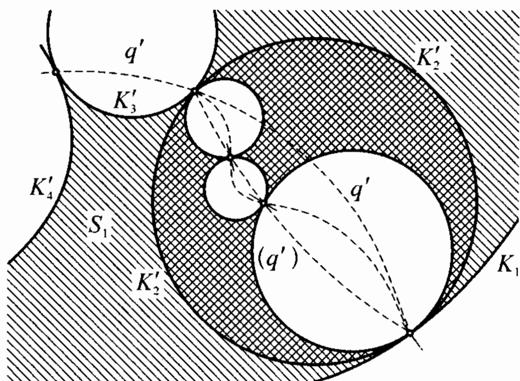


图 28.18

① 这 3 个新圆和 K_2' 一道，形成原先所取的两个圆弧四边形对 K_2' 的反演像。——中译者

圆内部又各得 3 个圆、两个四边形以及一个经过切点的圆弧。总之，圆弧(q')又被 3 个新圆弧构成的链所形成的曲线代替，它——和相切圆的中心线相继地交于$+\varphi_0$ 和$-\varphi_0$ 角——也没有折点。图 28.18 表示出这个链。

这里用 3 个圆弧代替截圆的一个弧的步骤，使我们想起在论述皮亚诺曲线时，由参数曲线 $\varphi_n(t)$和 $\psi_n(t)$分别作成 $\varphi_{n+1}(t)$和$\psi_{n+1}(t)$的情景(图 28.19)。现在回到把两个初始四边形对 K_1，K_2，K_3，K_4 反演所得的整个图像。我们得到 $4 \cdot 3=12$ 个新圆，4 个原有切点和 8 个新切点，这一共 12 个新旧切点都在一条由圆弧 q'，q''，q'''，q^{IV} 所构成的闭的截曲线上，把这样的反演施于每组 3 个新圆，就得到 $4 \cdot 3^2$ 条新的圆弧。它们相继接成一个由边界圆构成的闭链，同时在已有的 12 个切点上，又添上两倍那么多的切点。这时所有已得的 36 个切点都在一条闭曲线上，它由 $4 \cdot 3$ 段圆弧接成，这些圆弧相接处没有折点，而且它和接点处的中心线交错地作$\pm\varphi_0$角。因此，把反演接连进行 n 次，其结果是：



图 28.19

(1) 由第 n 次到第$n+1$ 次反演得到 $4 \cdot 3^{n+1}$个新的边界圆，穿过它们有一条新的截曲线；

(2) 共有 $4 \cdot 3^{n+1}$个切点；

(3) $4 \cdot 3^n$ 段圆弧，它们构成新截曲线。

你们可以理解，由此将得到什么。若令截曲线为 C_0，C_1，C_2 等等，则问题是：这个截曲线序列的极限图形 C_∞是什么样的？所有切点以及它们的聚点的轨迹 C_∞是否就是把内四边形和外四边形所产生的两个区网分隔开来的点集？

§ 28.5　所得非解析曲线的性质

要把 C_∞叫做一条曲线，我们还必须证明某些事实。但无论如

何，这个称谓是正确的。我们即将证明下面几个定理：

(1) C_∞是一条若当曲线，即C_∞可以用$x=\varphi(t)$，$y=\psi(t)$这样的形式表示，其中φ，ψ具有以前所规定的性质；

(2) 在它上面，切点处处稠密，而在这无穷多个切点，它的切线都是对应的中心线；

(3) 可是，只要像我们明确假定了的，φ_0不等于0，这条曲线就是非解析的。

有这些定理，我们先证明第二个。

为了证明定理，我结合以前曾经讨论过的一个论点，但在这里，联系着我们的特殊问题，我把它在较广泛的意义上复述一下。

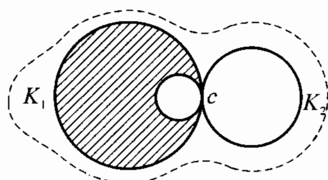


图 28.20

设c为我们将要考虑的点，我选两个在点c相切的圆，它们把平面分割成一个外部和一个内部(图 28.20)。

现在，我们设想如此作出整个点集C_∞：设想它落在两个相切圆外部那部分已经画出(我们还不知道按我们定义的约定，这个点集是不是一条曲线，这个问题和所要论述的无关)，然后通过多次反演，把所得图形表现在K_1和K_2的内部，使c成为极限点。问题是：我们能谈论该点集在点c的切线吗？

在下面情况下，就可以说，切线是存在的：设q是该点集的任意点，若令q无限制地靠近c，不管它靠近方式如何，弦$\overline{qc}$总靠近一个极限位置。

现在我们要说的是：无论怎样选取q，在外面那个区的某处，总有C_∞上一点p，这个点p通过对K_1的反演S_1，或通过对K_2的反演S_2，或通过一系列这些反演的结合，要变成q，即p可以看做q的等价点。因此，对应于弦$\overline{qc}$就有弦$\overline{pc}$。由此可见，若作出一切的弦$\overline{pc}$，并令p接连对K_1和K_2作反演，就得到弦$\overline{qc}$。我们先不谈论当

q——它总是属于点集 C_∞——无限制地靠近 c 时，弦$\overline{qc}$的极限位置，而谈论：当反演 S_1 和 S_2 无限制地作用于 p 时，弦$\overline{pc}$中的任意一条的极限位置。

这样就回到第 165—166 页讨论过的问题。在那里，我们看到，正是上面所说的极限步骤，使弦$\overline{pc}$以中心线为极限，这适用于每条弦$\overline{pc}$；一切弦$\overline{pc}$都以中心线为极限。现在我们可以指出，对于 C_∞ 中每一个趋于 c 的序列 q_n，弦$\overline{q_n c}$的极限位置总是中心线①。这样，中心线就可以看做是点集 C_∞ 的切线。

当我们考虑曲线序列 C_1，C_2，…时，这个结果的意义很大，因为这些曲线和在点 c 的中心线交错作 $\pm\varphi_0$ 角，而我们知道，其极限曲线 C_∞ 却和中心线相切。

为了使这种情况让人感到可以接受，我们举一个与此类似的例子。经过点 c 作一个序列的正弦曲线，如图 28.21 所示，它们的振幅是$\dfrac{1}{2^n}$，波长是$\dfrac{\lambda}{2^n}$，在 c 的位相是 $n\pi$（$n = 0, 1, 2, \cdots$）。

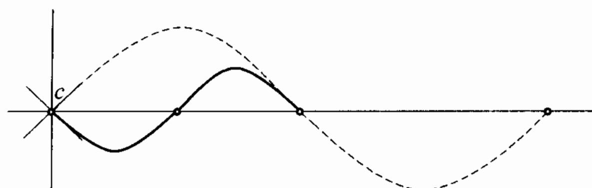


图 28.21

① ［对于序列 q_n 的每一点，有一个和 K_1，K_2 都正交的圆，这个点和它的所有等价点都在正交圆上。现在，经过 c，和 K_1，K_2 正交，而含有 C_∞的点的圆中，有两个属于最外面的。其中一个 H_1 经过 K_1 和 K_4 的切点，另一个 H_2 经过 K_2 和 K_3 的切点。理由如下：考虑由 K_1，K_2，K_3，K_4 所围成的四边形（参看图 28.14）。若对这个四边形的顶点无限制地施行变换 S_1，S_2，S_3，S_4，就得到在 C_∞上这些圆内部的所有切点，但这样，一个顶点在 K_1（或 K_2）内的像只能在 H_1 上和 H_2 上或在它们之间。切点的聚点也必然如此。因此，在 K_1 和 K_2 内部，曲线 C_∞ 只能夹在 H_1 和 H_2 之间，所以任意弦序列$\overline{q_n c}$（$q_n \to c$）可以放在两个弦序列$\overline{p_n c}$和$\overline{p_n' c}$之间，其中 p_n 和 p_n'依次是在 H_1 上和 H_2 上 q_n 的等价点］。

这些波形相似的正弦曲线 C_1，C_2，…在点 c 的切线和 x 轴交错地作交角 $\pm\varphi_0$。问题是，其极限曲线的切线是什么？对这个问题我们显然要仔细考虑曲线序列的极限。各个单个的辅助曲线 C_1，C_2，…的切线并不趋于一个极限位置（它们总是上下跳跃着），可是这些曲线本身却有一条极限曲线，即 x 轴，它在 c 有确定的切线，而这切线当然是 x 轴本身。

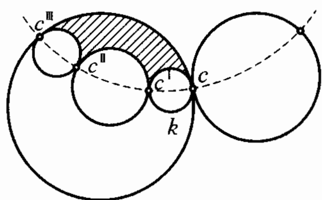


图 28.22

现在对于切点 c 及其聚点所构成的集合 C_∞，我们知道，在每点 c 都有确定的切线。但 C_∞ 却决不是一条解析曲线。说明如下：在我们图里的一系列圆中，仍取两个在点 c 相切的圆。在其中一个的内部，作两个圆弧四边形（图 28.22）。令这里出现的切点为 c，c^{I}，c^{II}，c^{III}。

我们知道，它们在一个截圆上，而且截圆在 c 和中心线作 φ_0 角，假定这个角是正的。

略去点 c^{III}，就有一个经过 c，c^{I}，c^{II} 的圆，它和 c 处的中心线作 $+\varphi_0$ 角。现在把两个四边形连同它们的截圆弧对图中用 k 标明的那个圆作反演。这样，在 k 里又得到 3 个新的圆，还有它们和 k 一起围成的四边形以及一段截圆弧。对经过 c 的那个新圆，把新的四边形及其截圆弧再作反演，并设想如此无限制地进行下去。这样，c^{I}，c^{II} 就无限制地向 c 靠近。我们得到点 c^{I}，c^{II} 所构成的序列，它们和 c 都属于点集 C_∞。每组 3 点 c，c^{I}，c^{II} 都在一个圆上，而这个圆则和 c 处的中心线作固定角 φ_0，φ_0 的符号则正负交替。于是得到点集 C_∞ 的一项性质，但我们即将证明，这个性质是解析曲线所不能有的，这个性质是：

和一点 c 无论多靠近处，有无穷多对点 c^{I}，c^{II}，经过它们和 c 的圆同 C_∞ 在 c 的切线作 $\pm\varphi_0$ 角。

对一条解析曲线，这的确是不能出现的。因为，若经过解析曲线上 3 点 c，c^{I}，c^{II} 作圆，并令 c^{I}，c^{II} 无限制地靠近 c，则这个圆就有确定的极限，即和曲线切线在 c 相切的密切圆。对于解析曲线，我们的圆绝不可能来回地和切线作有穷角 $\pm\varphi_0$。因此，曲线 C_∞，假如它是曲线的话，在任何点都没有密切圆。

还需要证明我们的点集 C_∞ 是一条若当曲线。

根据第 149 页的说明，若一个点集 C_∞ 可以一对一而且双向连续地映射到一个圆周上，它就叫做一条闭若当曲线。

对我们的点集 C_∞，只需把图 28.12 和图 28.17 相对照，就不难得到这样一个双向映射。

其法如下：令一个图中的初始四边形的顶点按循环次序和另一个图中的初始四边形的顶点相对应，与此相联系，再令两图中的旁四边形的顶点也如此相对应。根据已阐明的规律，一个图的切点显然和另一个图的切点也是一对一而且双连续地相对应。由于切点的集在各自有关的整体图形(即图 28.12 中的正交圆和 C_∞)上处处稠密，上述对应关系可以推广到有关的整体图形上：图 28.12 的正交圆上每一个非切点 α 可以看做一个切点序列 α_n 的极限点，令 α 同 α_n 在 C_∞ 上的对应序列的极限点相对应。这样，图 28.17 的点集 C_∞ 就成为图 28.12 中的正交圆的一对一而双向连续的映像，证明于是完成。据此，点集 C_∞ 的确是一条若当曲线①。

在结束关于自守图形的讨论时，在一定意义上，我已达到本讲演中关于精确几何部分的一个高峰，因此，看来可以在此对已获得的结

---

① [H · 庞加莱首先研究了这种非解析曲线(*Acta Mathematica* 第 3 卷，1883 年，第 77—80 页＝*Oeuvres*，第二卷，第 285—287 页)，克莱因在信中曾向庞加莱指出这种曲线存在。克莱因致庞加莱的这封信印在克莱因的数学文集第 3 卷，第 590—593 页上。可以参考该卷第 582 页上克莱因关于自守函数的史前史的叙述。

弗里克较深入地研究了，在上面所讨论的无正交圆的零角圆弧四边形的款中，通过反演以得到非解析的极限图形，*Math. Ann.* 第 44 卷，1894 年，第 565—599 页；还可参考弗里克与克莱因，*Theorie der automorphen Funktionen* (《自守函数》，第 1 卷，第 415—428 页)。]

果加以概括了。

和自守函数相联系，我排除算术步骤，用一种纯几何方法作出一个确定的无处稠密而又完备的点集，然后把其中初始图形加以特殊化①，以获得一条非解析曲线。这就表明，这些现代概念也能通过纯几何途径出现，因而并非只是分析家们的虚构。我把我的信念表达如下：

只要人们把精确数学的课题充分深入下去，集合论中的现代问题就会到处呈现出来。人们不得不对它们进行探究。

对于几何（以及一切涉及空间观念的数学业务），存在着下面两方面的作法；对此，人们往往进行得片面或者搞不清楚如何对待：

(1) 对近似数学按其本来面目来理解和对待；

(2) 另一方面，毫不躲避任何（在精确数学意义下）的理想化。

我要说："人们应当实行第一条而不放弃第二条。"

我要说明这个论点是如何适用于理论力学的，这是除几何外利用空间观点最多的科学。

这里，我们有以下问题：

在理论力学里，什么是近似数学？什么是精确数学？在力学里，这两方面在多大程度上相互推动，在多大程度上混淆不清？

力学一分为二：

(a) 第一种叙述力学的方式是和观察相结合，不越出经验领域之外。它所处理的不是质点而是小的个体之类。在自然科学中，这种只是描述直接观察到的事物而不去建立在现象后面的理论的倾向，人们一般称之为现象学。当我使用这个词时，显然是把现象力学②看做一种近似数学。具体细节从以前的解释可以推知（把现象限于用线性项表述这种做法不在此列，因为那样对现象的说明已是足够准确地近似了，如此等等）。

---

① 我们假定初始各圆循环相切。

② phànomenologische Mechanik. ——中译者

(b) 与此并列的有理想力学，这是我对它的称谓。这是以感知为基础建立数学概念，看看以合理的方法能由此引出什么。在这里，我们确实有质点以及这些质点相互作用中的严格规律，确实有导数等等。这种理想力学属于精确数学。由于人们对理想化毫不畏缩，就能讨论最广泛类型的点集如非解析曲线之类。

在上面提到的分野之外，对力学还有另一种流行的数学处理方式，那是(a)和(b)的混合物。它之所以流行，是因为在 18 世纪时，近似数学和精确数学的区别还没有显露出来，这种区别只是在 19 世纪才逐渐弄清楚的。因此，一切函数被看做本该是解析的，一切极限过程的次序都是可交换的（"那像在极乐世界那样"——P · 杜 · 布瓦-雷蒙）。

我的意见一贯是，对于初学者讲课总要采取这种一定程度上非逻辑的观点；但另一方面又要强调，当听讲者已获得较成熟的理解能力时，不能回避要划分精确数学和近似数学的已知界限。在力学教学中，也是如此。与此有关，我谈一谈《数学科学百科全书》中关于力学的论述。在那里，目前所说的现代观点（代表我个人的）只能附带提到。因为我的同事需要列出已有的文献，却不能把各科观点都写进去；而《百科全书》的目的只是收集资料，使以后的探索者不至于像经常看到的那样，因为不了解已有的成果而受到限制。

§ 28.6　这整个论述的前提，韦龙尼斯的进一步理想化

与此相联系，我必须指出，有些数学家要把理想化比我们所理解的再推进一步。在这里，我回到精确数学，并提出一本书，其出现曾经引起很大轰动：

Veronese, G: *Fondamenti di geometria*（韦龙尼斯：《几何学

基础》)，Padua，1891 年。由舍普(A. Schepp)译成德文，莱比锡，1894 年。

关于韦龙尼斯的基本观点，我可以作如下介绍①。我们一贯认为，精确几何的一切论点都建立在现代实数观念以及直线上的点与实数的一一对应关系的基础之上。人们是否采取这个基本态度，那本身不是数学问题而是是否适宜的问题。数学(狭义地)是从这个(或者别的)公理出发的。韦龙尼斯不是这样。

他设想有记号 η_1，η_2，…代表着实在的、阶数递升的无穷小量，对它们并且可以按一定方式施行运算，他还设想构造如下的式子：

$$x = \alpha + \alpha_1 \eta_1 + \alpha_2 \eta_2 + \cdots$$

其中 α，α_1，…是普通实数，然后把它界定为数轴上的点。因此，对他来说，在数轴上，我们的"有理点"和"无理点"是不够的，还要用无穷小量 η_1，η_2，…界定的点插进去。

这里要提出的第一个问题是：人们能否对这种式子作运算，而不出现矛盾？实际上这是可能的，因而从抽象数学的观点，不能反对韦龙尼斯的作法。但人们可以进一步问：承认在数学上可以允许，是否也就意味着宜于采用？

这样就涉及一个一般性问题：在数学中，根本上什么样的问题才是适宜的？

理论上，我当然可以提出任何问题，没人能禁止我去研究它。但只有涉及下面的问题才是适宜的：这些问题联系着人们反正必须研究而且由于涉及事物本质，也已经着手研究的其他问题。可以指出，在这个提法所划定的领域中，已有足够多需要完成的数学问题，要求用我们实际上非常有限的能力的一大部分去探索，把一

---

① 参考第一卷第 234—236 页(中译本第 249—251 页)和第二卷第 220—224 页(中译本第 242—246 页)关于非阿基米德数的论述。

些仍然不清楚的新问题整理好。也许我还可以这样说：当人们在研究一个新课题[①]中有必要深化理论认识时，进行探索是应当肯定的；若仅仅是为了求新而扩大研究，那是不值得的。

我还想指出，G·康托说过："科学的本质在于它的自由。"换句话说，当数学从其前提总能推得正确结论时，就可以自由自在地做下去。我理论上承认康托的说法，但同时我觉得非常重要的是，在实践中要补充一种制约：那就是，凡是倡议自由的人，都要承担一种责任。因此，在建立数学概念时，我不同意绝对的随意性，要对它用数学科学整体的观点来加以鉴别。

我就此结束离开本题的简短讨论，并转入近似数学领域。

① [为了避免误解，可以明确强调指出，非阿基米德数就是这样的课题(参看第一卷和第二卷的讨论)]。

第二十九章

转入应用几何：A. 测量学

§29.1 一切实际度量的不准确性，斯涅尼奥斯课题的实践

上面那一部分，我们专门讨论精确几何。现在进入另一方面，转到应用几何，即确实用到几何操作的几何。

我们把应用操作分为两类：度量和作图（包括模型制造），与此相应，就把应用几何分为两类：

(A) 测量学（有关度量的研究）；

(B) 作图几何（最广义的画法几何）。

我们的主要问题是：在这两个领域里，近似数学和精确数学是如何划分的？

对此，我回答如下：

测量学是几何学的一部分。在其中，近似数学的思路发挥得最清楚而且最彻底。在那里，人们不断地探究观测的准确性以及由观测所推得的结论的准确性。

另一方面，对于作图几何，还缺乏在近似数学意义上的合理发展。人们习惯地认为，不准确性是当然要渗进来的，把不准确性减小到最低限度是不能回避的课题。于是有如下规律：

“尽可能画得准确，但尽可能少地信赖其结果”（芬斯特瓦尔德（Finsterwalder））。

我先谈测量学，而且从初级测量学谈起，即从三角形和多边形的

度量谈起。我们要考虑度量的理论处理在近似数学思想中起什么作用。

一切度量分为两类，即长度和角度的度量，它们总带着不准确性。这有一系列的根源，略举如下：

首先是所使用的工具的有限准确性，因为它的刻度只准确到一定限度，即(长度)1 米或(角度)1 分的某个小数位。其次，这些刻度也受到外界的种种干扰(温度的变化，由于负荷所产生的弯曲、合金的化学变化等)。第三，还有观察者主观方面的干扰，即所谓的“个人方程”。最后，度量的对象也不是精确地给定的，因为经过大气的折射，光线线路并不是直线。

由于有这类讨厌的情况，度量是不准确的。至于度量的误差究竟有多大，就要根据情况作具体分析。我们首先要问，由(有限度准确性的)数据能作怎样的计算以得到别的数值？特别是，所计算出的数值将有多大的准确度？

对此，纯理论探讨也能起作用。作为典型的例子，可以讨论斯涅尼奥斯(Snellius)4 项课题①。

我可以简短地把它说明如下：设已经给定 3 点 A，B，C(例如海边 3 个航行标志)；问题是要确定另一点 P(海上的船只)的位置，方法是通过从 P 对 A，B，C 的观测，即测量 $\varphi=\angle APB$ 和 $\psi=\angle BPC$(图 29.1)。在理论上，问题是：若假定两个角的测量不准确度相等，由它们确定的 P 的准确度如何？

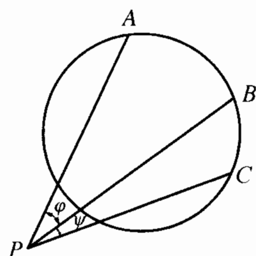


图 29.1

这里有一种重要情况，它产生在点 P 位于由 A，B，C 3 点所确定的圆上的时候。这时，无论点 P 在圆上哪个位置，φ，ψ 两角的值

① [往往错误地称为波特诺特(Pothenot)问题。]

都是常数。

在这种情况下，φ 和 ψ 两个角根本不能确定 P 的位置，它可以在那个圆上任意处。我们说，对于斯涅尼奥斯课题，有一个危险的圆，即经过 A, B, C 3 点的圆。

如果你们考虑到有这样一个产生不确定性的圆，你们就会了解，对于靠近这圆的点 P，测量两个角 φ, ψ 中的微小误差可能产生极为严重的后果。对于靠近该圆的点，φ, ψ 两角和对于圆上的点差别很小。因而测量它们中的微小误差，对于点 P 的确定，可能有十分严重的影响。所以，我们有如下结论：在危险圆的邻近的点，微小的测量误差一般地对点的理论位置要产生异常大的影响。

当点 P 在危险圆邻近时，确定其位置的准确性通常是很糟的；当它远离那个圆时，结果就会有本质的改善。

实际上，我们可以用“同等准确度曲线”把平面覆盖起来：在这样一条曲线上的两点 P 和 P' 由测到的 φ, ψ 值计算出来的位置有相等的准确度。根据 W · 若当的研究，这些等准确度曲线的形状就像图 29.2 那样。由这个图可以看出，在危险圆邻近，准确度的确很小，但在 3 个顶点邻近，有些地方准确度却很大①。

当然，斯涅尼奥斯课题中的上述问题在测量学的每个课题中都会出现。在这里，φ 和 ψ 两个角的测量能准确到什么程度，还都是未确定的因素。

① 参看 W · 若当：*Handbuch der Vermessungskunde*（《测量法手册》），第三版，第 1 卷，第五章。Stuttgart，1888 年。[图 29.2 采自若当的论文：“Über die Genaujgkeit einfacher geodàtischer Operationen”（“关于简单测量中的准确度”，*Zeitschr. f. Math. u. Physik* 第 16 卷，1871 年），第 397—425]。

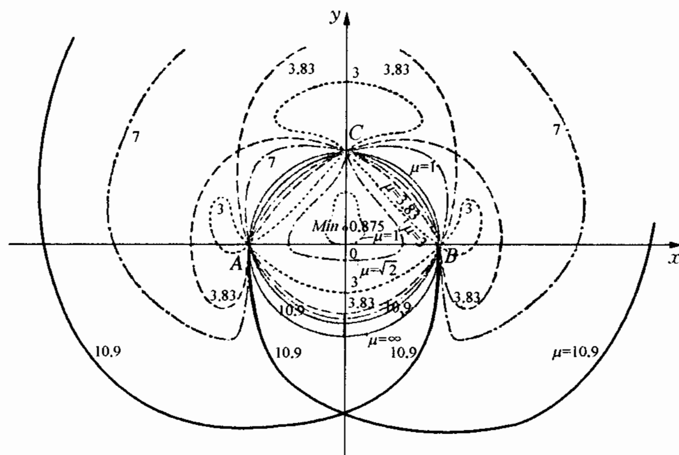


图 29.2

图 29.2 观测计算中的等准确度曲线。ABC 是一个等腰三角形。设 r 是危险圆的半径，M 是通过观测数据计算出来的点 P 位置的误差①中值，δ 是测量 φ，ψ 两角时的误差中值，则 $M=\mu r\delta$，其中 μ 满足方程。

$$r^4\mu^2=\frac{r^2+x^2+y^2}{(r^2-x^2-y^2)^2}(r^2+x^2+y^2-2ry)\cdot[(r^2+x^2+y^2)^2-4r^2x^2],$$

令 $\mu=$常数，则这个方程确定一条等准确度曲线，在它上面的点，其 M 值相同。

§ 29.2　通过多余的度量来确定准确度，最小二乘法的原则阐述

现在，如何从这样的准确度作出判断呢？这是另一个重要问题。我的答案是：

为了判断我们度量的不准确度，采用的法宝是平差法②，像最小二乘法所教导的那样。

情况是这样的：实际进行的度量项目多于确定结论的需要，然

① ［若对于点 P，P^* 是计算出的不准确位置，则距离 P^*P 就是这里所指的误差］。

② Ausgleichungsrechnung. ——中译者

后考察，这些受控制下的度量彼此吻合到什么样的程度。从它们合拍的性质和情况，人们就获得某种概率上的根据，据以判断那些单个的度量的准确度如何，以及所得结果应当在怎样的误差范围之内。我再次通过一个确定的例子来阐明这个问题。这个例子是：在一个已给底边上，确定一个四边形（这里不考虑底边的准确度）。我们可以通过在 A 和 B 的各两个角来确定 C，D（图 29.3）。但当着手用仪器来确定 C 和 D 时，我们添上 4 个角 γ_1，γ_2，δ_1，δ_2，也就是用 8 个角来代替必须的 4 个。我们首先要问，这 8 个角之间必须满足几个几何条件？当然，必须有 4 个条件。但是，当我们把所测得的数据代入时，这些条件只是带着某些误差地得到满足。我们说，8 个角度的度量含有 8 项“误差”，而且通常采取的步骤是，把 8 个测量数据加以修正。第一，使理论上的条件得到准确的满足；第二，使测量数据的改变量的平方和最小。

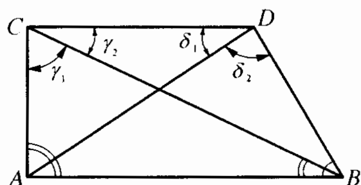


图 29.3

属于概率论的这个最小二乘法的根据是什么，已超出了我们讨论的范围；我们来直接谈它的应用。假定运用上述方法已经确定了 C 和 D 两点的位置。这些是它们的真正位置吗？如果是，那所用的就将是一种奇妙的方法，能从不好的观察，推得正确的结果。不能这样来理解最小二乘法。事实是这样的：在已经得到 C 和 D“最可能”的位置后，还必须用最小二乘法来考察这些位置的准确度（在这里，准确度的定义当然还有一定的随意性，人们谈到了误差中值、概率误差等等）。其结果是有分别含 C，D 两点在内的两个椭圆域，它们的大小同真实点在域内的概率有关。所以，还有对运用最小二乘法所得结果作出估计的技术问题。你们看，这涉及的就是某个近似数学课题的一个系统解答。问题根本不是要准确地确定 C，D 两点，而是要确定它们的大致位置以及这种不确定性存在于什么区域内。最

小二乘法对这个属于近似数学的课题作了确切的处理。

§ 29.3 近似计算，用关于球面小三角形的勒让德定理来说明

在处理这类近似问题中，要用到数值计算，作为又一个要点，我对数值计算的性质作些说明。当人们度量 1 米长度只需要准确到 $\frac{1}{100}$ 毫米时，就没有理由去给出 7 位小数的准确数字结果。因此，我提出的第一个要求是：在进行运算时，应当总是采用简约算法，即所用数值的位数只限于对有关问题有意义的范围内。

关于有时遇到的计算的简化，有球面三角里的一个有趣的例子，即“勒让德(Legendre)定理”。

它涉及一个“小”球面三角形和平面上与之对应的、具有相等的边长 a, b, c 的三角形之间的关系(图 29.4)。若球面三角形的角是 α, β, γ，而平面三角形的角是 α', β', γ'，勒让德定理指出，这两个三角形对应角之差是球面角盈的三分之一：

$$\alpha' = \alpha - \frac{\varepsilon}{3},\ \beta' = \beta - \frac{\varepsilon}{3},\ \gamma' = \gamma - \frac{\varepsilon}{3},$$

其中 $\varepsilon = \alpha + \beta + \gamma - 180^\circ$。

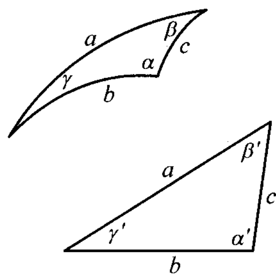


图 29.4

证明方法是利用级数展开，只保留其中次数最低的项。对此，我们不谈了，也不谈其误差估计[①]。无论如何，你们可以看出，这个定

① [其证明可参阅 E. Hammer: *Sphärische Trigonometric* (哈默尔：《球面三角法》，第四版，第 542—548 页，Stuttgart 1916 年)。在那本书的第 687—688 页上记述了关于这个定理历史的有趣说明。]

理提供多大的便利。运用平面三角公式比球面三角公式要容易得多[①]。我归结如下：

勒让德定理使数值计算容易得多；我把它作为一个典型而优美的例子，以说明什么是近似数学。

最后，可以谈一些历史过程：在测量学中，大体在 20 世纪初就普遍使用了最小二乘法、简约计算和勒让德定理等等。特殊地，高斯给出了这些方法的完整论述；我在这里不愿错过引用 1899 年豪克 (Hauck) 在门兴的自然科学家大会 (Münchener Naturforscherversammlung) 上所说的一句话：

“每一个土地测量者都涂上了高斯的一滴油。”

这意味着，每一个测量学家都要思考一下，什么是近似数学，他又应如何以适当方式运用数学来驾驭经验材料。

大约在 1860 年以后，纯粹数学就把高斯那滴油丢失了。那时，雅可比、高斯、泊松、柯西都已去世。此后，纯粹数学回避了那些似乎是低一等的东西。从这个角度来看，上述勒让德定理的历史提供了一个极好的事例。勒让德最初是在他的几何课本中给出他的定理的。但在这课本的新版中，它却被删掉了，尽管对球面三角法还是有着详尽的论述[②]。

§ 29.4 地球参考椭面上最短线在测量学中的意义（附关于微分方程论的假设）

关于粗浅的测量学就谈这些。我现在转到高级测量学。

---

① [高斯在他的 *Disquisitiones generales circa superficies curvas* (《关于曲面上曲线的一般论述》, Göttingen, 1827 年)中的第 24—28 篇论文 = *Ges. Werke* (《论文集》)第四卷，第 251—258 页里把勒让德定理推广到任意曲面上的短程三角形。]

② 在 19 世纪纯粹数学家们那里，近似数学的意义消失的另一个典型事例是，分析学的现代课本对早先人们曾经加以高度评价的差分方程几乎根本不再提了。只是最近在这里才开始出现转变(参看第一卷第 254 页，中译本第 271—272 页)。

在这里，地球表面是作为椭面[①]看待的。首先要考虑的是椭面上的短程线。我们要问，地球表面在多大程度上对应于一个椭面？而且，尽管地球表面有着无穷尽的不规则性（山、谷、树林、草地），在它上面，椭面上的短程线思想是如何能够保持的？

我这样说，好像是毫无必要的，因为每个人都认为这是理所当然的，而实际工作者也都是不自觉地这样看的。但仍有必要把事情搞得很明确，免得人们有时会使用精确数学的论点，而对于应用数学，这是不够的。

首先，理想曲面上的短程线 $x(t)$，$y(t)$，$z(t)$的精确定义是什么？

我们知道有两种定义：

一种方式是考虑曲面上连接两点的曲线弧长

$$\int\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2+\left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2}\,\mathrm{d}t$$

并要求它有极小值。另一方式是考虑曲线的密切面，要求在曲线每点，密切面和曲面切面垂直[②]。两种方式都用到微积分，而我们知道，微积分是建立在严格的极限概念基础上的。

在短程线的精确定义中，必须有精确界定的理想曲面，而$\frac{\mathrm{d}x}{\mathrm{d}t}$，$\frac{\mathrm{d}y}{\mathrm{d}t}$，$\frac{\mathrm{d}z}{\mathrm{d}t}$则表示差商的极限值。

在回顾了这些之后，我们把地球表面和一个椭面相比较。

像测量学家已经做的那样，人们当然可以在宏观上把地球表面和一个椭面相比较。我们可以参考伯尔施（A. Börsch）[③]的关于垂线

---

① 即地球参考椭面（简称参考椭面）。——中译者

② ［参看 W. Blaschke, *Differentialgeometrie Bd.* Ⅰ（布拉施克：《微分几何学》第1卷，第二版 1924 年，§ 56。）］

③ *Berichte über die Lotabweichungsbestimmungen*（《关于垂线偏差测定的报告》，国际大地测量会议。斯图加特，1898 年，1903 年）。

偏差测定的报告,其中讨论了被认为测量得最好的欧洲中部和一个椭圆盘的对比。但是,"微观上"它们根本没有一致性。例如,要在地球表面作一个切面,按照精密数学精神,就需要考虑到表面上任意小的变化,但是,那里根本不存在光滑的曲面,它的外表是极度不规则的。我们从看得见的微小变化(山与谷的差别、地面的不规则性、植被的变化等等),就已经看出事情是办不到的。当我们根据严格定义的要求,再深一步考虑到有机物的细胞构成以及物质的分子结构时,就更是如此。那里根本谈不上一般规律,谈不上一个严格的曲面,而只有一个具有多层次的空间,其复杂性和其宏观表现构成鲜明对比。我们必须说:当我们把地球和一个椭面对比时,根本谈不上地球表面的"细微处",谈不上小的差分和差商,更不要说给出微商(导数)了。

为了谈得略微生动些,我再迈进一步:

在实践中,什么是地球表面,并没有精确定义,因为并没有确切指出,要把哪一部分作为测量对象。不过,当我们在小范围内取差商时,差商的值总极度地来回摆动。你们看,在椭面上总不会近似地这样。在很小的局部,两者并不一致。但是,我要指出,当我们规定微商中出现的差分属于一定数量级时,它们却近似地一致。

设想有一块平坦地域(山地的讨论要困难些),在那里,我们可以确切地看到大约 10 千米远的地平线。平行于水平线的视线所构成的平面可以和椭面切面对比。更一般地,我们说:为了在大范围内考察地形(在这个范围内可以联系到椭面上的短程线的确定),并引进 Δx, Δy, Δz,以及比值$\frac{\Delta y}{\Delta x}$, $\frac{\Delta z}{\Delta x}$时,我们设想 Δx, Δy, Δz(或者更恰当地$\sqrt{\Delta x^2+\Delta y^2+\Delta z^2}$)的长度大约是在 10 千米之谱。这样所得的差商可以看做近似于椭面上的差商,就像利用观察到的水平线所得平面可以看做近似于椭面切面那样。对于一块山岭地带,我们当然要假定先把山和谷作一番平整工作,才能得出类似的关系。这

种设想，和我早先关于已经画好的曲线所说的是合拍的：在那里，当谈到那样的曲线的切线时，我们说，不能简单地通过微商（或者说根本不能通过微商，因为它根本不存在）来确定，而只能通过差商，其中的差分还是要属于一定数量级的。

对这些已经有了共同的认识后，我们要问，椭面上的短程线对实际测量起什么作用？显然不能把对前者所知的细节全部搬过来，并据此把椭面上的短程线和地球表面的最短线对比。下面我倒是要向你们说明测量家们实际采用的方法。

除了使用“短程线”这个名词之外，我要使用“短程多边形”这个词。一个已给曲面的短程多边形的边是具已给数量级长度的直线段，线段的端点在曲面上；此外，多边形还是“短程的”。这里所谓“短程”是指：或者在两个端点之间，它是这类多边形中最短的一条，或者在多边形的顶点，它的密切面①和曲面切面分别垂直（图 29.5）。我们首先设想，在代表地球的参考椭面上有这样一个短程多边形。这时我们选取其边长为 10 千米。这样就得到完全确定的一个多边形，它属于纯粹数学领域，因为它的定义是以完全确定的假定为基础的。另一方面，在实际地球表面可以毫无困难地作一条短程多边形，各边之长为 10 千米，其准确度则当然要按情况而定。我们只需从一个顶点向另一个瞄准，但要注意，从同一点出发的两条瞄准线所确定的平面②总垂直于通过地平面界定的切面。

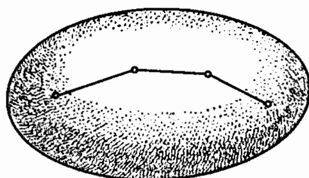


图 29.5

现在，在测量家们看来，这样实际作出的多边形和椭面上的理论

① 多边形的边都是直线段，而经过直线的每个平面都可以看做直线的密切面。——中译者

② 对于理想曲面，当多边形的边数无限制地增加，而边长都趋于 0 时，多边形的极限是曲面上一条曲线，这个平面的极限位置就是曲线的密切面。——中译者

多边形是近似地一致的，而这无疑是正确的。对此，我们始终留给测量家们去作出判断。他们无形中假定了纯粹数学中的一个定理，即理想椭面上的多边形是可以看做近似于曲面上的短程线的。这是近似数学的一个定理，它在我们所讨论的领域里得到检验。因此，我要说：尽管纯粹数学已为测量家们准备好了他们所需的知识，但只是通过现代微积分的极限方法来界定代表地球表面的理想椭面上的短程线是不够的。人们还要估计一下，在它上面从 A 到 B 的短程线，同它上面连接这两点，边长为 10 千米的短程多边形之间的差异有多大。于是就提出了这样的问题：在文献中，这些是否讨论了？在这方面，理论成就和应用需要之间是否也有令人感到遗憾的鸿沟？对此，可以这样回答，问题是讨论了的，但不是按我们感兴趣的观点来处理的。这问题应当按这里的精神加以阐述。

我把问题提得更一般化些，我们专门考虑通过微分方程界定的曲线。为了表述方便，我限于考虑含两个变数的一阶微分方程 $y'=\varphi(x, y)$。按照柯西的思路，这个微分方程解的存在性可以如下推得：取满足该方程的一系列元素 x, y, y'，使它们构成多边形，然后令多边形边数增加，边长缩小，以得到微分方程的一条积分曲线。下面我试行在另一个意义上，而且不那么明显地，运用柯西的思路来说明问题。

柯西把曲线和多边形作比较，求出它们之间的差异的估计，以推证积分曲线的存在。我们则利用这个估计来确认我们(在上述的特款中的)作法的合理性：把实际操作中所得到的地球表面上的短程多边形，不是用参考椭面上的理论短程多边形来近似地表示，而是用它上面的短程线！一般地，我要说：关于这个特款所说的话原则上也适用于运用(以精确的极限概念为基础的)微积分来处理经验世界中事物的一切情况。或者，从另一种角度来说：上述这类讨论对于保持现代形式的纯粹数学(其中微积分是以极限概念为基础的)和应用的联系是必要的。其间的纽带是近几年被忽略的、作为近似数学

一部分的、关于差分学的一个合理的论述。

§29.5　关于水准面及其实际测定

我顺便谈谈所谓的水准曲面(简称水准面)。

水准面是以牛顿引力理论为根据所引申出的地球重力场的等势曲面。

我们有势能

$$V=\int\frac{\rho \mathrm{d}k}{r}+\frac{\omega^2}{2}(x^2+y^2),$$

其中积分范围是地球的一切质量元 $\rho \mathrm{d}k$。第二项来自地球绕其轴(假定与 z 轴重合)自转的角速度 ω(=离心力势能)，V=常数，它确定表观重力[①](即离心力和实际重力之和)的等势面。令

$$X=\frac{\partial V}{\partial x},\ Y=\frac{\partial V}{\partial y},\ Z=\frac{\partial V}{\partial z},$$

就得表观重力的分力。等势曲面 V=常数，就称为"大地水准面"(Geoidflächen，这个名词来源于希腊文的地球，如 Geometrie=几何，Geographie=地理，等等)。显然，对应于不同的常数值，有一族水准面，其中经过某个选定点的称为主水准面。对于德国，选的定点在柏林天文台上某个固定点下 37 米处，和北海、东海水平面相差几个厘米。问题是：水准面是什么样的？人们怎样得到它们？对它们有哪些了解？

在这里，我要介绍关于测量的一些新结果，它使以前的假设有了修改。

从前假定了地球内部物质是均匀分布的。据此，计算了可见的

① 表观重力=scheinbare schwere。——中译者

地球的部分质量,并加上海洋中水的力量,这样,在大陆和山岭上的地域,所得的水准曲面非常向上突出。人们计算到地球上 200—400 米的高度。后来对重力强度的测量表明,所根据的假设是完全错误的,发现了大陆和山岭下面质量一般有欠缺,因而抵消了外部突出部分的质量。人们当然不会设想下面会有空洞,而是设想下面有质量较轻的岩石,像是在可见地面下有着处于静力平衡状态的液态下层。

在多种多样具体修正中,这是普遍的结果。无论如何,在新探究中发现,水准面比原先所期待的要平整得多。不过,有时和理论椭面相差许多米,可是这些恰恰是在人们比较没有预测到的地区。自然,对很大部分的地球表面还缺乏决定性的测量①。

下面我谈谈人们所掌握的确定水准面的实际方法。我特别要介绍。

(a) 天文定位法。它给出经过每点的等势面或水准面的法方向,从而给出其切面。

(b) 利用摆确定重力法。重力自然确定于它的分量 X, Y, Z。要得到水准面法方向的总力,就需要求 $P=\frac{\partial V}{\partial n}$,其中 n 是在有关点的法方向。在测量重力时,我们就是要测得 P 值,目的在于找出,在所考察的地方,等势曲面 $V=vc$ ($v=0, 1, \cdots$) 有多密。在重力值小的地方,水准面就较密(P 和 $\mathrm{d}n$ 成反比)。

(c) 第三是直接测量法,即在不同点作水平方向和垂直方向的测量。这个方法(在理论上)确定在各地所测得的水准面的不同点,以得到曲面的位置。

---

① [为了解这里所论的内容及文献,除了 F. R. Helmert, *Die mathematischen und physikalischen Theorien der höheren Geodäsie* (黑尔默:《高级测量学的数学理论与物理理论》两卷,莱比锡,1880 年和 1884 年)以外,皮策蒂(P. Pizzetti)在《百科全书》Ⅵ1,3 上的 *Höhere Geodäsie* (《高级测量学》,1906 年结束),黑尔默在《百科全书》Ⅵ 1, 7 的 *Die Schwerkraft und die Massenverteilung der Erde* (《重力与地球质量分布》,1910 年结束)是有帮助的。黑尔默: *Methoden der angewandten Geophysik* (《应用地球物理方法》)上面也列有许多文献,德累斯顿与莱比锡,1926 年]。

假定观测已经完备，我们就不但能得到水准面的点的相对位置，还得到它在各点的切面和势能沿法方向的导数。这里的数学问题是，在作了尽可能准确的一切测定之后，尽可能准确地确定水准面。我所说的前提是：一切测定尽可能准确；可这个前提，除了个别地域外，当然还并未得到满足。因此，所需的经验材料只是在地球表面一小部分是齐备的。我根本无法进一步谈论各种具体的观测，那是属于地球物理学所要阐述的事，我只强调一项突出的成果，那就是南森(F. Nansen)在(1893—1896 年)极地考察中所进行的重力观测，他发现法向重力和预期的一致。

但我现在要按照这个讲演的精神作出抽象数学的评论，我们问：

从根本上来说，水准面是精确地界定了的吗？

你们对于我下面的论点将不会感到惊奇：因为人们不能把水准面在大范围内看成确切的，更不要说是解析的曲面；这个论点我将加以阐明：

大地水准面的定义显然在理论上也是不完备的，像一切实际事物那样，它只是近似的。

对此我不作过于琐碎的分析，只作下面的论述：

首先，势能 V 是作为对于地球一切质量元的积分出现的。于是我们可以问：应不应该把大气也包括在内，等等。这些都只能按习惯加以规定；例如，我们可以把从地面上 10 千米内固定的、呈流体及气体状态的空气的质量算在地球质量之内。这样就把更高处的空气质量排除在外了；因为空气实际上延伸得很远，尽管很稀薄以至逐渐消失，但这些对于积分还是有影响的。此外，有些物质(水的环流、气压变化等)又是处于运动状态，因而势能 V 是时间的函数。当然，时间变化的影响不大，在实际工作中可以忽略，在理论考虑上却必须注意。

在这里，我特别要从我的观点谈论 H・布龙斯(Bruns)的一本书 *Die Figur der Erde* (《地球的图像》，柏林，1878 年)。这本书的特点是论述的数学严密性。我愿意问：他的论述在多大程度上和我

在这里的讲演中反复强调的精神一致？

布龙斯设想，地球是由某些不同的层构成的，如岩石、水，等等，像图 29.6 中所显示的那样。他假设这些物质互相连接，各层的密度分别是常数 S_1，S_2，S_3，…关于这个假设，如我们曾经谈论过的，最好还是采用平均值为好。布龙斯还假设相邻层的交接曲面都是“解析曲面”。这和我所持的、在自然界中根本不存在“解析曲面”的观点又是抵触的，因而只能有保留地去理解。所以，假设地球是不同的均匀部分所构成，而且各部分又以解析曲面为界，这决不是客观实际的准确描述，它只能是近似的。

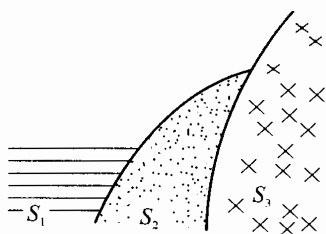


图 29.6

在上述假设的前提下，布龙斯推得了他理想的势能理论中的公式和定理，并提出问题：根据这个精确的势能理论，水准面的面貌如何？他得到的结论是，对于同一种物质的每个水准面，都表现为解析曲面片，而在两种物质相接处，两片解析曲面就无折点地相接，但在相接处曲率半径却不连续；这是因为，在理想势能理论中，由一种介质到另一种介质时，一阶导数是连续的，二阶导数则不连续。

我们要问：布龙斯的成果，或者任何类似的对实际度量的理论推断，其作用何在？因为严格地说，那些假设与实际不符。或者说，只能是实际情况的理想描述。联系到短程线对大地测量所起的作用，我要断言：研究布龙斯成果的意义，类似于上面所讨论的，研究椭面上的短程线的实际意义。

谈论其中细节将使我们深入到势能理论中去，因而只能就此打住，这样就结束了关于测量学的讨论。回顾我们的讨论，可以说：

测量学是一个光辉事例，它说明数学能对应用起什么作用以及如何起作用。当然，一切所得到的都只是近似的，而在探讨进行到最后时，总还要确定结果的近似程度。

第三十章

续论应用几何：B. 作图几何

§30.1 关于作图几何中一种误差理论的假设，用帕斯卡定理的作图说明

在应用几何的第二个分支，即作图几何，那里，事情远远不那么好。

作图几何的目的，可以是通过作图来表现空间关系（画法几何），也可以是在图纸上作图来代替数值运算（图解算法）。在两种情况下，以前所考虑的问题仍然适用：其准确度如何？所涉及的课题解决到什么程度？还有哪些东西尚待处理？

首先我要指出：

除了我将要谈到的个别开端性工作外，在作图几何里，至今还没有建立像测量学那样合理的误差理论。

所谓一个误差理论合理，是说，它是以使用概率的方法为基础的；为了判断一项作图方法的准确度，人们对同一个课题重复作图，然后用最小二乘法或其他办法把所得结果加以调整。

作为这样一种合理误差理论的建设性解答的一次初始尝试，可以介绍法国数学家莱莫恩（Lemoine）开始的工作，它称为几何图解法。对于一项作图课题，例如阿波罗尼斯问题（用圆规和直尺作8个圆与3个固定圆相切），可以有不同方法，莱莫恩把各种方法所使用圆规和直尺的次数作为衡量该作图法的简单性或复杂性的尺度。你们可以看到，这和我们要考虑的问题有什么联系。这个数目越大，即圆规和直尺用得越多，一般来说，作图结果就越不准

确。因此,我们乐意把莱莫恩意义上的最简单作法看做最准确的作法。可是,作为衡量"简单度"的数字只是个很表面的尺度,因而这种认定只能是初始性的。原因是,那里把不可比的东西作为可比的看待了。例如非常靠近的两点的连线比相距较远的两点的连线要作得较不准确[①]。

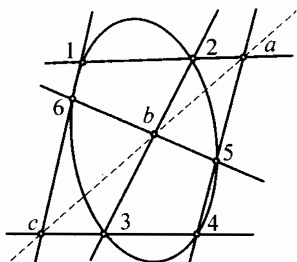


图 30.1

我用一个有代表性的例子来说明。设通过作图来重现帕斯卡六角形或有关的帕斯卡定理。

这定理是：设在一条二次曲线上有 1, 2, 3, 4, 5, 6 共 6 个点,作连线 $\frac{\overline{12}}{\overline{45}}$ $\frac{\overline{23}}{\overline{56}}$ $\frac{\overline{34}}{\overline{61}}$,则对边交点 a, b, c 在一条直线上(图 30.1)。

这个定理可以在通常公理的基础上严格证明。与此平行,现在我要提出一个广义的帕斯卡定理[②]。它暂时可以叙述为:"设有大致在一条二次曲线上的 6 点,作大致的连线并取其交点 a, b, c,则这 3 点大致在一条直线上。"当然,主要问题是,把"大致"用准确而适当给定的节来代替。

显然这个定理(而不是"理想的"帕斯卡定理)是作图时的真正理论基础。

证明这个定理可以完全采用解析方法,我们早已说明,这是最方便的方法。我们先采用解析方法来表达帕斯卡定理本身,即我们设想,通过解析方法进行如下:

---

① 参看 A. Witting: *Geometrische Konstruktionen, insbesondere in begrenzter Ebene. Programm des Gymnasiums zum heiligen Kreuz* (维廷:《几何作图,特殊地在有限平面内。关于神圣十字架的中学计划》),德累斯顿,1899 年,维廷给出了辅助作图以避免不准确的步骤,如避免用直线连接两个靠近的点。此外,再参考 P. Zühlke: *Konstruktionen in begrenzter Ebene* (齐尔克:《在有限平面内作图法》,莱比锡,1913 年)。

② 我简短地说这定理是"近似的",因为它涉及的是近似问题。若限制它的幅度,它本身就是准确的(就像近似数学中每个完备的,即确定了误差幅度的定理那样)。

设二次曲线上 6 个点 1，2，…，6 有坐标(x_1，y_1)；(x_2，y_2)；…；(x_6，y_6)，而由它们导出的 3 个点 a，b，c 有坐标(x_a，y_a)；…；(x_c，y_c)，则表明 a，b，c 在一条直线上的关系是

$$\begin{vmatrix} x_a & y_a & 1 \\ x_b & y_b & 1 \\ x_c & y_c & 1 \end{vmatrix} = 0。$$

现在把坐标(x_1，y_1)；…；(x_6，y_6)改为 $x_1+\delta x_1$，$y_1+\delta y_1$；…；$x_6+\delta x_6$，$y_6+\delta y_6$，其中 δx_v，δy_v 相对于图中的其余数量是小的。为了不使事情弄得过分复杂，我们假定，此外不再出现不准确性，即假定连线是准确地画成的，交点也是准确地得出的。于是交点 a，b，c 的坐标的变化显然只与 δx_1，δy_1；…有关。我们对于变化了的坐标计算行列式

$$\begin{vmatrix} x_a+\delta x_a & y_a+\delta y_a & 1 \\ x_b+\delta x_b & y_b+\delta y_b & 1 \\ x_c+\delta x_c & y_c+\delta y_c & 1 \end{vmatrix} = \Delta,$$

它的几何意义是三角形面积的两倍。这个 Δ 当然不再等于 0；要考察的是，在适当地选取 δx_v，δy_v 的情况下，它是否保持有小值。

在测量学和度量天文学中，这种步骤是通常采用的。在那里，这种公式称为微分公式，于是在这里，我们的要求就可以说成：我们不但要有帕斯卡定理公式，还要有相应的微分公式。这样帕斯卡定理的作图验证才有了充分理论基础[或者说得更好些，才有对作图验证的数学理解①].

① [和莱莫恩的几何图解有着不同基础的一种关于几何作图准确性理论，可以参考 F. Geuer，*Die Genauigkeit geometrischer Zeichnungen*（戈伊尔：《几何作图的准确性》）(Jahresbericht 1902 des Progymnasiums in Durlach in Baden). P. Böhmer，*Über geometrische Approximationen*（伯默尔：《关于几何近似》）(Diss. Göttingen，1904). K. Nitz，*Anwendungen der Theorie der Fehler in der Ebene auf Konstruktionen mit Zirkel* （转下页）

我还要对“微分公式①”这个词略微阐述，以免使用它时有所疑虑。

其所以会有疑虑，是由于一方面它涉及以严格的极限概念为基础的微分运算；另一方面，在使用微分公式时，又涉及小的差分。这个矛盾可解决如下。

（接上页）*und Lineal*（尼茨：《平面上的误差理论对圆规直尺作图的应用》）（Diss. Königsberg 1905）和“Beiträge zu einer Fehlertheorie der geometrischen Konstruktionen”（“关于几何作图的一种误差理论的报告”）（*Zschr. f.* Math. u. Phys. 第 53 卷，1906 年，第 1—37 页）。此外，关于误差的讨论还值得参考 *H. Schwerdt*, Lehrbuch der Nomographie（施韦尔特：《列线图解教程》，柏林，1924 年）。最后，还可以参阅第 11 页所提到的范伦的书中“几何图解与误差理论”一章，第 121—129 页。

戈伊尔工作的基本思路是运用高斯的调整方法。每个求作的点或每条求作的直线都作多次，然后加以修整；使作图时误差平方成为最小。在修整中的修改量就作为评价作图准确度的依据。伯默尔采取的方法是书中所推荐的微分公式的办法。例如在他的论文里，人们看到已给三边求作三角形的问题。他用有关的微分公式来判断先作哪一边，再作哪一边，以达到尽可能高的准确度，对于帕斯卡定理，他也用了与此相应的方法，即用了微分公式，但不就此结束，他的论文是从解答一个修整课题导出的。那课题是，已给大致在一个二次曲线上的 6 点，求“最近”的二次曲线，他所谓“最近”是指满足彭塞列-切比雪夫修整要求的二次曲线，即要求所给点到二次曲线的最大距离为最小。但下面他要讨论相应的高斯修整要求。

尼茨论文的主要成果利用了测量家所建立的平面上的误差理论、他的思路可以概括如下。用笔画出的直线总是有宽度的。若画两条相交直线，则其相交处实际是在一个平行四边形里。若这两个笔画宽度相等，则平行四边形是菱形，假定高斯法则适用，则当我们用圆规尖端来寻求两线交点时，具有相同概率的点在一个椭圆上，椭圆中心在平行四边形对角线交点，而平行四边形的边的方向对于椭圆为共轭方向，令概率从 0 到 1 之间变化，就得到一族共轴而相似的椭圆。其中有一个，称为“平均误差椭圆”就取作“把圆规尖放在两线交点上”的操作准确度的表征，若用图纸上两个小圆来代替两点，而用直尺作连线，则人们发现，具相同概率的连线包络一个双曲线。若令概率从 0 到 1 变化，就得到一族共轴而相似的双曲线，其中有一个“平均误差双曲线”。它用来衡量“依靠两点作连线”的操作的准确度。

一切圆规直尺作图都可以由 5 种基本操作构成。除了已指出的两种外，还有以下 3 种：(a) 把圆规尖放在一条直线和一个圆的交点上；(b) 把圆规尖放在两个圆的交点上；(c) 以已给点为中心，作出已给半径的圆。在讨论这类作图的误差时，就要考虑另外 3 条平均误差曲线，现在，对于任意一项圆规直尺作图，若要确定其平均误差，就首先要利用其中“点”的直径和“直笔”的宽度的平均误差来确定所用到的每项基本操作的误差曲线。在此基础上，再应用误差综合理论，从基本操作中的平均误差计算整个作图的平均误差。

尼茨通过这个方式研究了一些简单作图，如线段的中垂线、直角、平行线等，并获得很有趣的结果。]

① [关于这个问题，瓦尔特在 1926 年的哥廷根假期课程中作了报告，其经过大为补充的讲稿即将印行。]

试取最简单的款：一元函数 $y=f(x)$。实际工作者是这样做的：为了得到 $f(x+\delta x)$，把它按泰勒定理展开，并截取其第一项，得

$$f(x+\delta x)=f(x)+\varphi(x)\delta x。$$

再令 $\varphi(x)=\frac{\mathrm{d}f}{\mathrm{d}x}$。我们当然要更仔细些，我们要说：$\varphi(x)\delta x$ 根本不是泰勒级数的第一项，它是余项，其中还必须加上一些已作为 0 处理的项；要把公式写得确切，$\varphi(x)$ 不能等于 $\frac{\mathrm{d}f(x)}{\mathrm{d}x}$，而必须等于 $\frac{\mathrm{d}f(x+\theta\delta x)}{\mathrm{d}x}$ $(0<\theta<1)$，即 $\varphi(x)$ 不是 $f(x)$ 在 x 处的导数而是在一个中间地方 $x+\theta\delta x$ 的导数。因此，这里用到 $\frac{\mathrm{d}f(x)}{\mathrm{d}x}$ 的一个中值。我们称公式

$$f(x+\delta x)=f(x)+\delta x\cdot f'(x+\theta\delta x)$$

为中值定理，于是可以说：

当我们在简化运算中，用 $f(x)+\varphi(x)\delta x$ 代替 $f(x+\delta x)$ 时，$\varphi(x)$ 不是由泰勒级数确定，而是按中值定理确定的。

对于复杂的款，微分公式的情况也和这个简单款一样。在应用中，人们往往把泰勒定理中增量里的高阶项略去。更准确地说，微分公式应当看做按中值定理规定的余项公式。当导数 $f'(x)$ 在 δx 里变化很小时，而且只有在这时，在应用中，前者和后者才没有实质的差别。

§ 30.2　由经验图形推导理想曲线性质的可能性

现在我想提出另一个，也是特别有趣的问题。

在第 153 页和以后几页中，我们看到，每一条画出的曲线可以用一条理想曲线代替，使得不但经验曲线的纵坐标，而且还有方向和曲率——只要它们能定量地确定——都为理想曲线的纵坐标、方向和曲率所满意地接近。现在问题是可否由我所看到的经验曲线显示出的关系推得关于理想曲线的相应的性质？

回答这个问题所采取的途径显然决定于上述的关系，即理想曲线是从感官直觉引出，只是通过定义而存在的。因此，我不能单纯依靠直观。相反，必须考虑的总是，由经验所作出的图像，呈现于我们眼中的概略事物能不能，或者为什么能，移植到理想图像上来，而移植时必须以有关定义为基础，并且用的是精确方法。

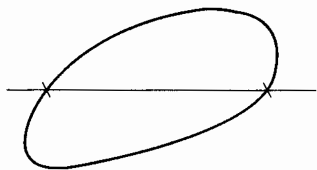


图 30.2

我们举一个例子。在图 30.2 里有一条处处是凸的闭曲线和一条与它相交的直线。我们注意到这里有两个而且只有两个交点。能把这作为定理移植到与之对应的理想图像吗？

我们设想用一条理想直线代替图中的直线，至于图中所画的曲线，我们首先用任意一条近似的若当曲线（它不一定有切线，更不用说有曲率圆）来代替。按照直观，要求理想直线既经过若当曲线的外部，也经过它的内部。因此，根据若当定理，在经验曲线和经验直线相交处附近，理想图像也确实有交点。但这个结论的依据是我们对若当曲线所下的定义。如果只为了达到和经验曲线近似的目的，我们可以不用若当曲线来代替经验曲线，而用一个处处稠密但不是闭的点集（即不包括其一切聚点的点集，如数轴上一切有理点所构成的集）。这时候，我们所说的交点就未必存在。

按上面所说，首先只是确定了交点的存在而没有确定交点的数目。还很可能，在经验图形只有一个交点处的邻近，理想图形却

有 3 个或 5 个交点（这个数目必是奇数，因为在那里，理想直线要从若当曲线外部进入内部）。这的确是不能排除的，因为还没有在定义上对那条若当曲线加以别的限制；我们只要想一想魏尔斯特拉斯曲线或皮亚诺曲线的状况就可知了。现在，我们选取的限制是，按照经验曲线的原型，理想曲线在每点应有确定的方向和曲率。此外，像经验曲线那样，它还应当是处处凸的，因而没有拐点。据此，我们就确实能证明，在经验曲线呈现出一个交点处，理想曲线也只有一个交点。

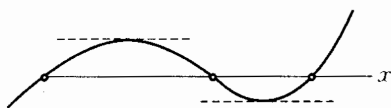


图 30.3

证明如下：假定有 3 个或更多交点，我们选取那条理想直线为 x 轴（图 30.3）。设曲线在这些交点间的一段的方程是 $y=f(x)$，其中函数 $f(x)$ 有连续的一阶和二阶导数。由于 $f(x)$ 在这段 x 轴有 3 个或更多零点，根据罗尔定理，$f''(x)$ 在该节中至少改变符号一次。这样，我们的理想曲线就将有拐点，这就和我们对它的假设矛盾。

于是，证明就完成了。经验图形为这证明提供了大方向，它对细节考虑也有启发作用；但最后还必须回到精确几何的概念和公设来证明。

在文献中，对这个问题的讨论有详有略，要看其作者兴趣是比较倾向于具体定理的严密基础还是于新结果的发现。两者各有道理，关于对具体问题有仔细阐述的作品，例如

Kneser, A.: "Einige allgemeine Sätze über die ein-fachsten Gestalten ebener Kurven"（克内泽尔："关于具最简单形状的平面曲线的若干一般定理"）。*Math. Ann.* 第 41 卷，1893 年，第 349—376 页。

另一方面，关于许多新定理的陈述有：

Juel, C.: "Einleitungin die Lehre von den graphischen Kurven"（丹麦文）（于尔："图像曲线理论导引"）*Mém. Acad. sc.*

Kopenhagen,第 6 辑,第 10 卷,1899 年,第 1—90 页①。

§30.3 对代数曲线的应用,将要用到的关于代数的知识

我本人愿意回到关于平面代数曲线实拐点的研究,若干年前,我曾谈论过这个问题,那时我充分使用了直观图形②。为此,对于所讨论的正则理想曲线就要再加一个条件,它是代数曲线,而且下面还要进一步限于讨论不具有"高阶"奇异性的 n 次曲线。

首先对 n 次平面代数曲线作简短论述。平面 n 次代数曲线 C_n 是用一个含 x, y 的 n 次方程③

$$Ax^n + Bx^{n-1}y + \cdots = 0$$

确定的。

若数一下里面有几项,再设想用一个不等于 0 的系数遍除各项,就可以看出,里面共有 $\frac{n(n+3)}{2}$ 个常数。

---

① [与此有关,可以举出于尔在自然研讨大会(斯图加特,1906 年)所作的很有意义的报告。它印在 *Jahresberichte der Deutschen Mathematikervereinigung* 第 16 卷(1907 年),第 196—204 页上。还可以参看于尔:"Einleitung in die Theorie der ebenen Elementarkurven dritter und vierter Ordnung"("平面三次和四次初等曲线导论")(*Mém. Acad. sc. Kopenhagen*,第 7 辑,第 9 卷,1914 年)和 *Math. Ann.* 第 76 卷,1915 年,第 343—353 页的论文,以及耶尔姆斯勒夫:(除第 17 页已列出的那篇"现实的几何"外)的 *Introduction á la théorie des suites monotones* (《单调序列理论引论》)(Bull. Acad. sc. Kopenhagen, 1914 年,第 1—74 页)和蒂默丁主编的 *Handbuchs der angewandten Mathematik*(《应用数学手册》)1914 年,第二部分的 Darstellende Geometrie(《画法几何》)第 135 页起。最后;关于卵形线理论的:Brunn, H.:Über Ovale und Eiflächen (布龙:《关于卵形线与卵形面》), Diss. München,1887 和 Exakte Grundlagen für eine Theorie der Ovale (《卵形线一个理论的严格基础》),Sitzungsber. der math.-phys. Kl. der Kgl. bayr. Akademie der Wissenschaften,第 24 卷(1894 年)第 93—111 页。]

② [克莱因:"Eine neue Relation zwischen den Singularitäten einer algebraischen Kurve"("代数曲线的奇点间一项新关系")。*Math. Ann.* 第 10 卷,1876 年,第 199—209 页,重印在 F·克莱因的《数学论文集》第 2 卷,第 78—88 页(1922 年)。]

③ 这是非齐次方程。——中译者

问题是 C_n 的形状。一条 C_n 由几条卵形线[①]或其他线路构成？它的奇点，特别是它的拐点情况如何？这些全面的问题足以构成代数曲线理论中的一大章；特别是对于低次曲线，人们已在这方面取得非常有趣的结果。我在此当然不能作系统的说明，我只限于谈论实拐点的数目。

我们把曲线方程简写成 $f(x, y) = 0$，或者，为了便于进行一般性讨论，采用射影几何方法，把它写成齐次方程

$$f(x_1, x_2, x_3) = 0，其中 x = \frac{x_1}{x_3}, y = \frac{x_2}{x_3}。$$

在采用齐次坐标考察拐点时，就用到由函数 $f(x_1, x_2, x_3)$ 的二阶偏导数构成的黑塞(Hesse)行列式

$$\Delta = \begin{vmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{vmatrix}。$$

当 f 是 n 次时，每个 f_{ik} 是 $n-2$ 次，因而 Δ 是 $3(n-2)$ 次。

根据贝祖定理，$f=0$ 和 $\Delta=0$ 有 $3n(n-2)$ 个交点；因此曲线 C_n 有 $3n(n-2)$ 个拐点 $[w = 3n(n-2)]$。这是为了确定代数曲线奇点的第一“普吕克公式”。(Plückersche Formel)更确切地，可以说：

由于曲线 $f = 0$ 和 $\Delta = 0$ 交于 $3n(n-2)$ 点，C_n 一般有 $3n(n-2)$ 个拐点。在这里，有待考察的是，有些交点可能重合，以及在 C_n 的这些交点中，有多少被可能有的奇点所吸收。

据此，对于低次的款：

$n = 2$ 时，$w = 0$；

$n = 3$ 时，$w = 9$；

$n = 4$ 时，$w = 24$。

① 这里的“卵形线”实际是指可能有拐点的闭曲线(参阅第 222 页底注)。——中译者

当我们画出一条 C_3（或 C_4）时，我们最多只能找到 3 个（或 8 个）实拐点，因而 G·萨尔蒙在他的 *Higher plane curves*（《高次平面曲线》）中提出猜想，在拐点中，至多有$\frac{1}{3}$，即 $n(n-2)$ 个是实的。这是我们准备弄清楚的问题。

需要证明的是，C_n 的实拐点数不能超过 $n(n-2)$，因而在最有利的情况下，至多只有三分之一的拐点是实的。

除了已给出的关于拐点的公式之外，我再不加证明地写下关于 C_n 二重切线数的普吕克公式：

$$t=\frac{n}{2}(n-2)(n^2-9)$$

（在此，再次假定了曲线没有所谓的奇点，我们即将回到这个问题）。

我们先讨论 C_n 的类数①，即由任意不在曲线上的点到曲线可能有的实切线和虚切线的数目。这些都是代数曲线理论中最初浅的事物。为了得到证明关于实拐点数的定理所需的公式，我必须对这个问题加以阐述。我采用较详尽的叙述方式。

先写出曲线在它上面一点 $x\equiv(x_1, x_2, x_3)$ 处的切线方程

$$\left(\frac{\partial f}{\partial x_1}\right)y_1+\left(\frac{\partial f}{\partial x_2}\right)y_2+\left(\frac{\partial f}{\partial x_3}\right)y_3=0,$$

其中 $y\equiv(y_1, y_2, y_3)$ 表示切线上的流动点的坐标，诸括弧表示导数在点 x 的值。在点 x 的切线方程，用动点 y 的坐标表示，可以简写成

$$f_1y_1+f_2y_2+f_3y_3=0。$$

但若令不在曲线上的点 y 固定，而把 x 看做在曲线上的动点，则上述方程显然确定一条 $(n-1)$ 次曲线；因为 f_1，f_2，f_3 对 x 是

① Klasse. ——中译者

$(n-1)$ 次的，它称为 y 点对于曲线 C_n 的第一极曲线（简称第一极线）。因此，这表明，若 C_n 在它上面的点 x 的切线经过 y，则 x 在这条 C_{n-1} 上，即它同时属于 C_n 和 y 对于 C_n 第一极线，因而是它们的交点。若能判断这样交点的数目，我们就知道 C_n 有几条切线经过 y。但根据贝祖定理，这个数目是 $n(n-1)$，或者说，C_n 的类数 k 是 $n(n-1)$。

这自然是确定切线的纯代数方法，它没有区别实切线和虚切线。我们还必须考虑到，曲线上可能有奇点，这时 k 的数目就要减小。因为不止一个 x 要落到这样的奇点上，而我们不愿意把这样的直线 yx 算作 C_n 的切线。

于是问题是：当奇点存在时，关于 k 的公式要如何变化？首先，根本什么是奇点？在这里，我们也只是给出最初浅的、我们所需用的东西，我们直截了当地界定：

若在一点 x，3 个导数 f_1，f_2，f_3 同时等于 0，则曲线在该处有奇点，这时切线方程为任意点 y 所满足。若二阶导数或二阶和三阶导数等都等于 0，或者它们满足某些条件，例如 $f_{11}f_{22}-f_{12}^2=0$，则曲线在 x 有高阶奇点。

为了了解奇点的情况，最简单的办法是把坐标三角形的一个顶点 $x_1=0$，$x_2=0$，$x_3=1$ 放在奇点（图 30.4）。这时曲线方程就简化如下：

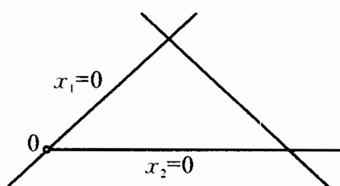


图 30.4

首先，每个 n 次方程可以按 x_3 的降幂式列为

$$f(x_1, x_2, x_3)=x_3^n\varphi_0+x_3^{n-1}\varphi_1+\cdots+x_3^0\varphi_n=0,$$

其中 φ_0，φ_1，…，φ_n 是 x_1，x_2 的 0 次，1 次，…，n 次齐次函数。由于 $x_1=0$，$x_2=0$ 在曲线上，常数 φ_0（零次多项式）等于 0，若按上述定义，这一点是奇点，则线性项 φ_1 也不出现，于是方程从二次式 φ_2 开始：

$$0=x_3^{n-2}\varphi_2+\cdots+x_3^0\varphi_n。$$

把 φ_2 写出来，例如

$$\varphi_2=a_{11}x_1^2+2a_{12}x_1x_2+a_{22}x_2^2。$$

一种特款是 φ_2 恒等于 0，另一种特款是 $\varphi_2=0$ 有重根 $\frac{x_1}{x_2}$。若不让曲线系数进一步特殊化，即假定在那里只是有奇点，则方程 $\varphi_2=0$ 给出两个不同的 $\frac{x_1}{x_2}$ 值。结果是我们的 n 次曲线 C_n 在原点有“二重点”，经过它，曲线有两个“分支”(图 30.5)。

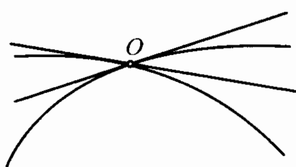


图 30.5

这样一个二重点对于曲线的类数 k 所起的作用是很容易看出的。

$y\equiv(y_1, y_2, y_3)$ 点的第一极线是

$$f_1y_1+f_2y_2+f_3y_3=0。$$

现在令 $f=x_3^{n-2}\varphi_2+\cdots+x_3^0\varphi_n$。我们立即看到，这条第一极线经过二重点 $x_1=0$, $x_2=0$ 一次，它的方向一般和 $\varphi_2=0$ 所确定的原曲线两支在二重点的切线都不相同。因此，二重点吸收了原曲线和第一极线的两个交点。于是有定理：

若原曲线有 d 个简单二重点，则曲线的类数 k 减小 $2d$，因为落在每个二重点上有原曲线和第一极线的两个交点，而我们不把二重点 x 和(不在曲线上的任意点)y 的连线看做 $f=0$ 的切线。

对我前面提出的问题，这个结论就是普吕克公式后面所要起的作用。现在我们转向普吕克处理过的两个问题：二重点吸收几个拐点？吸收几条二重切线？请让我指出以下一点。

设想把具有任意系数的 f 确实地写出来，同样写出方程 $f_1=0$, $f_2=0$, $f_3=0$(根据已知的一个欧拉定理，由这 3 个方程即得 $f=$

0)。这 3 个方程是否相容？或者说，f 的系数要满足什么条件，奇点就存在？

对于一个 C_2 的款，这是容易写出的，设

$$f = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3 + a_{33}x_3^2。$$

这时由 $f_1 = 0$，$f_2 = 0$，$f_3 = 0$，有

$$\frac{1}{2}f_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = 0,$$

$$\frac{1}{2}f_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = 0,$$

$$\frac{1}{2}f_3 = a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = 0。$$

它们是 3 个含 x_1，x_2，x_3 的齐次方程，只有当条件

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = 0$$

得到满足时，它们才是相容的。于是有熟知的定理：

若二次曲线有奇点，则系数行列式必等于 0。

对于高次曲线，情况也一样。这时，$f_1 = 0$，$f_2 = 0$，$f_3 = 0$ 给出 3 个 $(n-1)$ 次方程，它们必须相容。于是消元式，或说原曲线的判别式 D 要等于 0，在 C_2 的特款，判别式等于行列式 Δ。沿用上面的说法，“一般地”，我说：

若一条 C_n 有奇点，则其系数的判别式必等于 0，因此，一般地（即只要其系数不满足上述条件）曲线没有奇点。

对于高阶奇点可以给出类似定理。联系着上面的讨论，我指出进一步的结论：若 C_n 有高阶奇点，则除 $D = 0$ 外，还要满足别的代数条件；因此，当 $D = 0$ 时，一般地，曲线在所考察的点(x_1, x_2, x_3)

只有简单的二重点，它使曲线的类数减小两个单位，要保证高阶奇点存在，就必须满足别的一些条件。

这就是一般的代数基础。但我们还没有对实与虚加以区别，连 $f(x_1, x_2, x_3)$ 的系数也都可以是虚数。

§30.4　提出所要证明的定理：$w' + 2t'' = n(n-2)$

我现在要考察的是，在一般代数理论的基础上可以获得什么样的实定理。我们首先要作一系列的准备。我们假定 $f(x_1, x_2, x_3)$ 的系数是实的。或者更适当地说，它们在实域里变化。这样，我们就一下子把目标放在一切 C_n 的整体上，而问题是，我们对 C_n 的形状能说些什么？

其次，我们把拐点数 w 分为实和虚两种，由于虚拐点必然成对出现，我们令 $w = w' + 2\bar{w}$。

同样，我们把二重切线也分为实和虚两种。其中实二重切线又分两种情况：一种是具有两个实切点的；另一种是具有两个虚切点的，这时我们看到的是一条孤立的二重切线。因此，在实二重切线中，设用 t' 表示具实切点的切线数，用 t'' 表示具虚切点（孤立二重切线）的切线数。于是二重切线总数 t 是

$$t = t' + t'' + 2\bar{t},$$

其中 $2\bar{t}$ 是虚二重切线数。

二重点的情况与此类似。我们可以写出

$$d = d' + d'' + 2\bar{d},$$

其中 $2\bar{d}$ 表示虚二重点数，d' 表示具实分支的二重点数，d'' 表示实孤立二重点数。

这一切不过是一般地说明若干记号将用来表示什么。现在我们

先要问，无论有没有二重点，一切曲线 C_n 的形状是什么样的。

我们先回顾一下 $n = 2$ 时曲线的形状。

若二重点不存在，就有一条实的或虚的曲线。我们要指出，椭圆、抛物线、双曲线 3 种实曲线将看成是没有本质区别的，因为每一种都可以经过射影变换化为另一种。若二重点存在，则 C_2 或者退化为一对实线或者一对虚线。在后一款，我们就看到一个孤立点。这种情况可以简单地由双曲线或椭圆经过极限过程产生。

椭圆、抛物线、双曲线称为偶性的，因为一条直线和它们或者交于两点，或者根本不相交。

现在考虑 C_3，它和一条直线总交于 3 点或一点，因此，总要出现一条所谓奇性的线路，像图 30.6 所画的（连同它的渐近线）那样。此外，还可能有一条偶性线路。

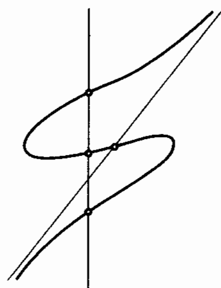


图 30.6

这种区别可以推广到 C_n。对于 C_n 的形状，可以带有某些不确定性地说：

没有奇点的 C_n 含有有尽多条闭线路。若 n 为偶数，则所有线路都是偶性的；若 n 为奇数，则除了可能有的偶性线路外，还有一条奇性线路（不会有两条或更多条奇性线路，因为那样它们将彼此相交，与无奇点的假设矛盾）。

至于其中的细节就要作特殊的分析或进一步深入的研究。好在对于我们将要讨论的关于拐点的定理，这些完全不必去了解。我们要证明的是：

对于无奇点的 C_n 在实拐点和孤立二重切线之间，有关系

$$w' + 2t'' = n(n-2)。$$

实际上，当 n 较大时，C_n 的形状是很多样性的，但这个关系对于各种各样无奇点的 C_n 都成立。

§ 30.5　证明中将采用的连续性方法

显然，若这个定理成立，则 $w' \leqslant n(n-2)$，因而萨尔蒙通过归纳所得到的关于实拐点的结论是正确的：C_n 根据普吕克第一公式，所有的拐点中，至多有 1/3 是实的。

为了证明我们的定理，需要说明一些辅助性的概念。

我们已经看到，C_n 的方程有 $\dfrac{n(n+3)}{2}$ 个常数，我们把它们看做高维空间的点的坐标，而把这个空间对应于 C_n 的点作为 C_n 的“代表点”。令 C_n 取一切各种可能的形状，即令 $f=0$ 的系数任意变化，则这些代表点充满整个 $\dfrac{n(n+1)}{2}$ 维空间。你们将要看到，利用这个空间来和一切 C_∞ 的整体相对照，将有多大的帮助。因为这使我们较易从事某些流形的考察。我们都主要地已经习惯于讨论点空间中的连续性问题。

现在，在所有的 C_n 中，考虑那些判别式 D 等于 0 的曲线，即至少有一个(可能更多)奇点的曲线。由于 $D=0$ 是关于 $f=0$ 的系数之间的一个代数方程(它是通过消元法得到的)，则 $\dfrac{n(n+3)}{2}$ 维代表空间，$D=0$ 代表一个 $\dfrac{n(n+3)}{2}-1$ 维流形。这个流形我们就称为曲面①，这是把 R_3 中熟悉的术语移用于 $R_{\frac{n(n+3)}{2}}$。

现在，在高维空间，这样的曲面形状是什么样的呢？从平面曲线的多样性可以设想，它会是很复杂的。但在这里，情况是简单的，因为我们所遇到的是代数图形。我立即可以指出：

一个代数曲面由空间有尽多块墙壁所构成，它把空间分隔成有尽

① 在高维空间，现在称为“超曲面”。——中译者

多个相遇于墙壁的区域；墙壁的维数比空间维数少一个。特殊地，这适用于我们的代表空间里的曲面 $D=0$。

现在，平面上判别式 D 等于 0 的 n 次曲线有哪些可能的情况呢？

首先是具有一个普通二重点的曲线，这个二重点使曲线的类数降低两个单位；其次是具有高级奇异性（这包括同时出现的多个二重点的情况）的曲线。但若存在着高级奇异性，则除 $D=0$ 外，还有其他代数条件方程要满足，不过这些条件只有有尽多个。由此可见，只要 n 有尽，就只有有尽多的不同情况：D 的根有些可能重合，或者不确定，等等。对于下面的讨论特别重要的是，这种条件只能有有尽多个。

那么，关于高级奇异性的这些条件方程对于其对应的代表点又有什么样的限制呢？我立即指出下面的定理：代表着具有高级奇异性的代数曲线的点，在曲面 $D=0$ 上构成至多有尽多的“代数曲线”，即维数比 $\frac{n(n+3)}{2}$ 小两个单位的代数流形。

在上面所说的可以用下面的空间图像来直观地表现：具有高级奇异性的代数曲线对于把空间分隔为区域的墙壁 $D=0$ 来说，提供装饰物，它们不会把墙壁布满（图 30.7）；它们本身的维数至多达到 $\frac{n(n+3)}{2}-2$。

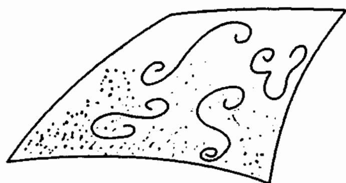


图 30.7

你们可以作下面的具体图像。取 R_3 作为墙壁，取一个立方体的 6 个平面，设想它们代表 $D=0$，而立方体的棱则对应于高级奇异性。这就是联系着下面定理的思想，该定理是后面讨论的基础。

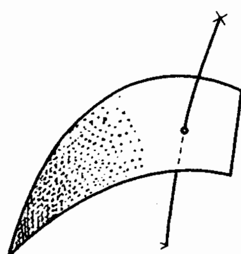


图 30.8

考虑一条无奇点的代数曲线，这就是在代表空间给定了不在曲面 $D=0$ 上的一点。令曲线系数连续变化，就可以从该点达到另一点。但要达到一个已给的终点，就可能必须穿过墙壁有尽多次(图 30.8)，但不和它上面的装饰物相遇。用另外的说法：

从任何一条无奇点的曲线，可以通过中间状态的曲线，以连续变化的方式，达到另外任何一条无奇点的曲线，这些中间曲线都没有高级奇异性，但其中至多有有尽多条具有一个普通二重点，而在这些二重点，曲线可以有实的或虚的分支。

这个定理以及它所根据的观点是我们论证的核心。我们要问：这个定理和我们早些时候的曲线概念相容吗？

如果我们所考虑的不是代数曲线①而是任意曲线，那么结论就不成立，例如皮亚诺曲线就能充满平面上一整块。可是，如果对曲线概念加以限制，则若当曲线已能把平面分隔为一个内部和一个外部。我们现在考虑的只是代数曲线和代数曲面，则(尽管值得做更仔细的分析)我们可以立刻作出下面的论断：

我们所阐述的高维空间的一般连通关系，即：由于考虑的是代数曲线和代数曲面，认为一个曲面起墙壁作用而曲线则不然，肯定是合乎实际的。

可是还有一个问题：为什么上面谈到墙壁把高维空间分隔成区域时，用了“有尽多个”这个词？为什么是“有尽多”？这是人们对其必然性可能提出的另一个疑问。对此，我可以说，若存在着无穷多的区域，则必存在着奇点的聚点，而这是和代数函数的性质不相容的。在这里，我不能谈其中细节，只是重复一下已经说过的：

① 原文作“曲面”。——中译者

我们所考虑的第一点是空间被分隔为区域，第二点是区域数是有尽的。

§ 30.6　有与无二重点的 C_n 之间的转化

在为以后讨论作了如此最一般几何的基础论述之后，我们问：

在实领域里，通过系数的连续变动，一条代数曲线是怎样由无二重点转化为有一个简单二重点（这时曲线的类数降低两个单位），然后又转化为无二重点的？

这里我先纯经验地作一个图形，对它进行考察，然后看如何能由此得到关于代数曲线的确切定理。我们按照第 205 页的思路来实现由经验曲线到理想曲线的过渡。

图 30.9 表示一条无二重点的曲线 1 的两支由上下两方趋近，融合出一个二重点，出现一条具二重点的曲线，然后曲线 2 又出现在左右两方。这个图的特点在代数曲线也能出现。我们的问题是要显示那些在极限过程中不能直观地掌握的变化，因为这种变化可以说是微观地实现的。

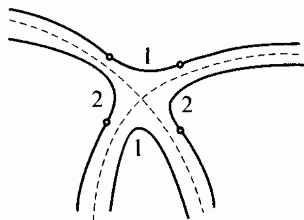


图 30.9

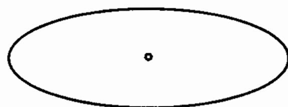


图 30.10

但这图只说明了两种可能性之一。第二种，同样可能出现的情况是，作为曲线组成部分的一条卵形线缩小成一个孤立二重点，然后消失，像图 30.10 所示那样。

现在，仍然在经验领域里，我们考察拐点。由图 30.9 我们发现

一个值得注意的现象：二重点在出现的时刻吸收了两个实拐点，然后又把它们释放出来。或者，换一种说法：曲线 1 和曲线 2 在二重点附近各有两个拐点，在极限过程中都落到二重点上了。——图 30.10 说明：

当一条卵形线缩成一个二重点然后消失时，如果事情是像图所表示的那样，则没有实拐点参与其中。

对于理想代数曲线，其对应情况如何呢？特殊地，和最后定理相联系：若一条代数曲线所含有的一条卵形线通过连续变形缩成一个孤立二重点，在卵形线接近转化时，它是否一般地没有实拐点？为什么？这里的"一般地"排除了 $f=0$ 的系数(除 $D=0$ 外)，还满足高级代数条件方程的情况。特殊地，那个二重点只使曲线的类数减小两个单位。我们需要考察，在讨论中的情况，能否有蜿蜒状的卵形线①缩成那个二重点(图 30.11)？

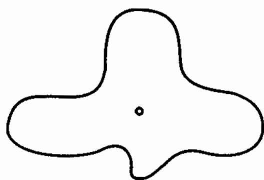


图 30.11

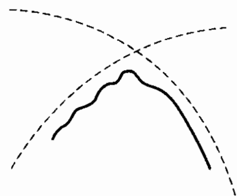


图 30.12

同样，对于具实分支的二重点，我们也可以问：在代数曲线上，除了图 30.9 所显示的必然存在的两个拐点外，是否会有更多拐点吸入二重点或者由二重点放出(图 30.12)。问题是，能否断定。一般地，这是不会出现的。我对卵形线的款进行关于这个问题的证明，因为这最为方便。

我假定卵形线有两个拐点。就有不在卵形线上的一点 p，从那里可对它作 4 条切线(图 30.13)。

① [在这里和以后，"卵形线"这个词表示偶性闭线，不一定是处处凸的。]

首先，可以作直线和卵形线交于 4 点，我们只需用直线把两个拐点连接起来，而正是这两个拐点把卵形线分成两段凸线。选取这条直线为 x 轴，同它垂直的一条直线为纵轴，则 x 轴把卵形线截成 4 段，在每一段上有纵坐标的一个极大值或极小值，卵形线在那里的切线都和 x 轴平行。因此，从 x 轴的无穷远点就可以对卵形线作至少 4 条实切线。把这个作法略微变动，就可以得到整个区域，从其中每点 p 可以对卵形线作至少 4 条切线。

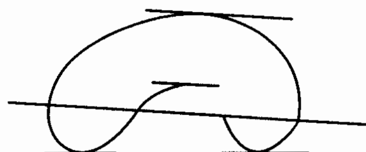


图 30.13

由此便可以引出矛盾。因为令卵形线缩成二重点 x（在取极限过程中始终保持 4 个拐点）时，上述那样的每点 p 到卵形线的 4 条切线显然都落到连线$\overline{xp}$上，因而对于这些点 p，二重点要吸收 4 条实切线。于是由代数观点，对于每一点 p，它要吸收 4 条切线，可是我们明确假定了，我们的二重点只是把曲线类数降低两个单位。由此可见，卵形线直到取极限前还有两个或更多个拐点这个假设是不能成立的。

与此相应，可以处理具有两条实分支的二重点(图 30.14)。

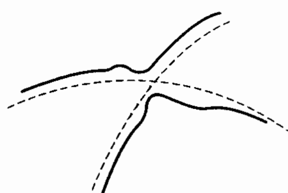


图 30.14

若具有两条实分支的二重点的近邻曲线(我用它称与具二重点曲线相邻的曲线)有多于两个必然存在的拐点，则当近邻曲线转化为具二重点的曲线时，曲线的类也要降低多于两个单位，仍然与我们明确假设转化时只降低两个单位矛盾。于是就证明了下面的定理：

若一条代数曲线通过连续变形转化为具一个二重点的曲线时，曲线的类数降低两个单位，则可以区别两款：在二重点的两个分支

是实的还是虚的，若是实的，则二重点吸收两个拐点；若是虚的，则二重点不吸收拐点。

从本讲演的观点来看，这个定理的推得，在一定程度上构成关于代数曲线拐点的考察的核心；因为它清楚地显示了两个方面的对比：一方面是突出直观；另一方面是根据定义的概念推导。

§ 30.7 符合定理的偶次曲线举例

我现在进入讨论的较简单的一节，面临的课题是构造出一些确实满足所要证明的关系 $w'+2t''=n(n+2)$ 的代数曲线。有了具体例子，我们就较易信服，这个关系是普遍成立的。我们将看到，通过曲线最一般的连续变形，这个关系保持不变。

构造这样的曲线，最简单的办法是：第一步，作出由低次曲线组成、具尽可能多的二重点的曲线，使人们可以掌握曲线的全貌；第二步，让这些曲线转化为无奇点的近邻曲线，转化后，我们要验证的关系仍然得到满足。

在这里，宜于分别考虑 n 为偶数与奇数两款。首先设 n 为偶数。我们先令 $n=4$，这时，我想向你们说明的事物容易用图显示。

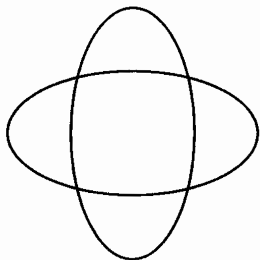


图 30.15

我们简单地作两个全等椭圆 $\Omega_1=0$ 和 $\Omega_2=0$，它们有共同的中心，其长轴则互相垂直，因而它们相交于 4 个实点(图 30.15)。取 $\Omega_1\Omega_2=0$ 为我们退化的 C_4。显然，它有 4 个二重点，即两椭圆的交点；而且由代数(根据贝祖定理)可知，它也没有其他二重点，于是曲线的二重点都是实的，而且曲线在那里都有实分支。

我们的曲线的类数是 $k=4$，因为每个椭圆的类数是 2。对于一般的 C_4，从不在曲线上的一点 y，可以作 $(4\cdot 3=)12$ 条切线。因此，

它的类数降低了 12－4＝8 个单位，由此可知，每个二重点使我们的 C_4 类数降低了两个单位。

其次，考察我们 C_4 的二重切线。根据一般定理，它的数目应是 $\frac{n}{2}(n-2)(n^2-9)$，故 $n=4$ 时，即是 28。在图中，这 28 条二重切线是：两个椭圆的 4 条公切线和 4 个二重点的 6 条连线，但这 6 条每条计算 4 次，其所以如此计数，可以按下述方法推出：设每条连线计算 α 次，则 6 条连线代表 6α 条二重切线，加上椭圆的 4 条公切线也都是二重切线，就得 $4+6\alpha=28$。下面我们从别的考虑来得出 4 这个因子，若再考虑到我们的特殊的 C_n 没有拐点，就我们的需要说，对 C_n 的了解就已全面了。现在我们考虑 C_4 的近邻曲线：令 $\Omega_1\Omega_2=\pm\epsilon$，其中 ϵ 是小的正量。

为了作具体分析，令

$$\Omega_1=\frac{x^2}{a^2}+\frac{y^2}{b^2}-1,\ \Omega_2=\frac{x^2}{b^2}+\frac{y^2}{a^2}-1。$$

这样，我们就看到，在椭圆 $\Omega_1=0$ 内部，Ω_1 小于 0，在它外部，Ω_1 大于 0，相应的情况对 $\Omega_2=0$ 也成立。因此，在一个椭圆内部而在另一个椭圆外部那 4 个区域，$\Omega_1\Omega_2$ 小于 0（图 30.15）。

现在先令 $\Omega_1\Omega_2=-\epsilon$，就得到完全在这些区域内无奇异点的一个近邻曲线，它表现为 4 个卵形线（图 30.16）。

你们看到，近邻曲线恰好有 8 个实拐点（$w'=8$），这是必然的，因为根据一般定理，每个实二重点应给出两个拐点而椭圆本来是没有拐点的。

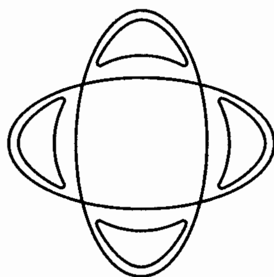


图 30.16

这个新作出的曲线的 28 条二重切线又如何呢？对当前的款，它们是容易全部核实的（图 30.17）。首先是两个椭圆的 4 条公切线直接变成我们的 4 个卵

形线的 4 条公切线，即 4 条实二重切线。其次是两个椭圆的交点的 6 条连线每条分解为 4 条实二重切线，如图中所示。因此，共有 28 条。我们的曲线不存在孤立二重切线。因为 28 条已作出的二重切线都有实切点，根据一般理论，不会有别的二重切线。于是 $t''=0$。因而确有

$$w'+t''=8=4\cdot 2=n(n-2)。$$

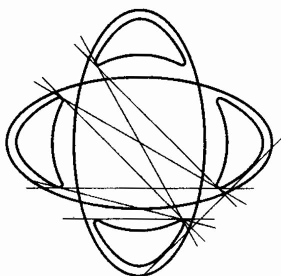


图 30.17

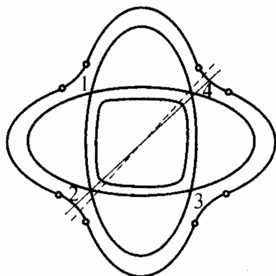


图 30.18

所以我们的第一个例子满足这个关系。

对于另一种近邻曲线 $\Omega_1\Omega_2=+\varepsilon$，情况有所不同，但结果是，$w'+2t''$ 的值仍然相同。

这时的曲线为两个分支构成，外部一条有 8 个拐点，内部一条没有拐点(图 30.18)。这样的曲线称为环形曲线：它的一个分支含在另一个分支的内部。

首先，原来的每个二重点仍然给出两个实拐点。其次，两个椭圆的公切线也仍然转化为二重切线，至于根据一般定理应当存在的其余 24 条二重切线则必然是虚的。例如，我们考虑原来两椭圆交点 2 和 4 的连线 $\overline{24}$，它应当产生 4 条二重切线。但我们看到，当它从原来位置略微变动时，它总是和曲线交于 4 点，它们每两个都不能会合。因此，若在 $\overline{24}$ 邻近有二重切线，它只能是虚的。于是有结论：

我们的环形曲线有 4 条实二重切线，每条有两个实切点，另有

24 条虚二重切线，但没有孤立实二重切线。

因此，用我们的记号，$t''=0$，关系 $w'+2t''=8$ 仍然成立。

这样，$n=4$ 的款就暂时解决了，这是指：$n=4$ 时，我们已经构造出满足我们关系的例子。为了对更高的偶值 n 作同样的事，我们先取 $\frac{n}{2}$ 个全等的椭圆，它们有同一个中心，前后两个的主轴成 $\frac{2\pi}{n}$ 角。例如，若像图 30.19 那样取 4 个椭圆，就得一个 C_8。对这样的 C_n，我们像前面对 C_4 那样作出各种计数。

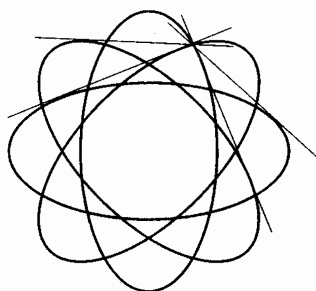


图 30.19

(1) 有多少二重点？

每两个椭圆交于 4 点，因而共有

$$\frac{4\frac{n}{2}\left(\frac{n}{2}-1\right)}{2}$$

个二重点：即曲线有 $\frac{n}{2}(n-2)$ 个实二重点，它们都有实分支。

(2) 我们计算曲线的类数 k。显然 $k=n$，这是因为从不在曲线上的一点可以对每个椭圆作两条切线，而我们共有 $\frac{n}{2}$ 个椭圆。若回到一般公式，则 $k=n(n-1)-R$；其中 R 是由于二重点存在而降低的数字，这个数字是

$$R=n(n-2)。$$

即由于奇点存在，类数比无奇点曲线降低 $n(n-2)$ 个单位，每个二重点降低两个单位。

(3) 我们来验证曲线有 $\frac{n}{2}(n-2)(n^2-9)$ 条二重切线。我们有

(a) 每两个椭圆有 4 条公切线，它们都是整个曲线的二重切线。这方面的二重切线是

$$\frac{\frac{n}{2}\left(\frac{n}{2}-1\right)\cdot 4}{2}=\frac{n(n-2)}{2}。$$

(b) 诸椭圆的相互交点的每条连线(不一定是同一对椭圆的交点的连线)给出 4 条二重切线。这给出

$$\frac{4\frac{n}{2}(n-2)\cdot\left[\frac{n}{2}(n-2)-1\right]}{2}$$

条二重切线。

(c) 最后,还有第三种二重曲线。它们是这样得来的:由两个椭圆的每个交点,可对其余 $\frac{n}{2}-2$ 个椭圆分别作两条切线。可以用不同的方法证明,作为整个曲线的二重切线,这些切线都要加倍计算。我们的椭圆相互有 $\frac{n}{2}(n-2)$ 个交点,从每个交点可以作 $\frac{n-4}{2}\cdot 2$ 条上述那样的实切线。因此,整个曲线上这种二重切线的数目是

$$\frac{n}{2}(n-2)\cdot\left(\frac{n}{2}-2\right)\cdot 2\cdot 2。$$

把 3 种二重曲线的数目相加,就得二重切线总数

$$\begin{aligned}&\frac{n}{2}(n-2)(1+n^2-2n+2n-2-8)\\=&\frac{n}{2}(n-2)(n^2-9)。\end{aligned}$$

于是从我们的特殊曲线,就验证了一般曲线应有的二重切线总数为 $\frac{n}{2}(n-2)(n^2-9)$。这些二重切线有一部分 4 条重合;另一部分两条重合,但都是实的,并且具有实切点。

(4) 我们的退化曲线当然没有拐点。

这个由$\frac{n}{2}$个椭圆$\Omega_1=0$，…，$\Omega_{\frac{n}{2}}=0$构成的退化n次曲线的方程是

$$\Omega_1\Omega_2\cdots\Omega_{\frac{n}{2}}=0。$$

由它出发，我们作近邻曲线

$$\Omega_1\Omega_2\cdots\Omega_{\frac{n}{2}}=\pm\varepsilon。$$

退化曲线的$\frac{n}{2}(n-2)$个二重点给出这曲线的$n(n-2)$个实拐点。

二重切线可以部分是实的，部分是虚的；这和转化到近邻曲线的方式有关，即和选取ε前面的符号正或负有关。但可以完全肯定的是，不会有孤立二重切线，因为由退化曲线的形状可知，只要二重切线是实的，它的切点也是实的。

于是我们已达到目的，我们有

$$w'=n(n-2),\ t''=0,$$

因而

$$w'+2t''=n(n-2)。$$

这就是所要验证的关系。

对于任意偶数n的例子是$n=4$的例子的推广，只是在计算二重切线时，多了(c)的一种。

§30.8　奇次曲线的例子

我现在转而考虑奇次n的款，我仍然从最低次，即从$n=5$入手，它对一般款有典型性。

和上面的讨论完全对应，我们从构造一个退化的C_5开始，把情

况弄清楚，然后上升到较一般的 C_5。于是我们要验证的关系是

$$w' + 2t'' = 15。$$

假定我们特殊的 C_5 为一条具方程 $\varphi = 0$ 的 C_3 和一个椭圆 C_2 ($\Omega = 0$) 组成，因而它的方程是

$$\varphi \cdot \Omega = 0。$$

问题是，如何选取 C_3 和 C_2 以便得到所需的关系？

我曾经指出，一条非退化的 C_3 总是表现为奇性线路。我们在此把它画成(图 30.20)只有一条渐近线。还可以把它画得对于渐近线对称，这时一个拐点落在曲线和渐近线的交点 O 上，因而其余两个拐点和 O 在一条直线上，分别在 O 的两侧，距离相等。通过压、伸和推移，即通过适当的仿射变换，可以使 C_3 具有像图 30.21 中那样的形状。我们不需要了解 C_3 的一切形状，只需取它有特别适合我们目的的形状。

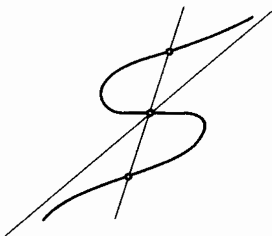


图 30.20

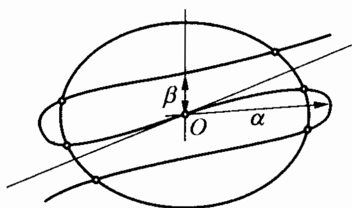


图 30.21

令 β 为从点 O 到曲线一侧的最短距离，而 α 为从它到一个顶点的距离(图 30.21)，并选取正数 a, b，使

$$a < \alpha,\ b > \beta, 但\ a > b。$$

以 O 为中心，以 a, b 为半轴作椭圆 C_2，则 C_3 和椭圆按照贝祖定理所可能有的 6 个交点都是实的。即使把椭圆绕 O 转动，也就是无论它的主轴如何改变方向，6 个交点都还是实的。另一方面，两个这样的椭圆相交于 4 个实点。我们将利用这些来适当地构造高次曲线

C_7, C_9,…这就不多谈了。

仍然回到我们的 C_5,对于它,问题完全可以具体地弄清楚。首先是它的类数。

显然 C_3 的类数是 $3 \cdot 2 = 6$,C_2 的是 $2 \cdot 1 = 2$,因而 C_5 的类数是 $6 + 2 = 8$。由于根据一般定理,$k_5 = n(n-1) = 5 \cdot 4 = 20$,这说明降低了 12 个单位,即每个二重点还是把类数降低两个单位。

实拐点数是 3,因为 C_3 有 3 个拐点,而椭圆没有。

一般的 C_n 有 $\frac{n}{2}(n-2)(n^2-9)$ 条二重切线,当 $n = 5$ 时,这数是 120。现在对我们的 C_5 加以验证。首先是 C_3 和 C_2 的公切线：由于 C_3 是 6 类的,C_2 是 2 类的,故这些公切线给出 C_5 的 12 条二重切线。它们是实的还是虚的?从图 30.21 可以看出,它们不会都是实的。对此,我们不深入探究,但可断言：C_3 和 C_2 的实二重切线都有实切点,而不会有共轭虚切点,因而不会有所谓的孤立二重切线。这是因为,这样公切线的两个切点应分别在 C_3 和 C_2 上,因而满足不同的代数条件。

其次,我们若从曲线 C_5 的一个二重点作直线,和曲线在另一点相切,则在二重点,这条直线总不会和椭圆 C_2 相切而只能和 C_3 相切。经过每个二重点,有 4 条这样的二重切线。这里的数字是这样得到的：从不在 C_3 上的任意点,可以作 6 条切线和 C_3 相切。当这个点趋于 C_3 上时,其中两条和 C_3 在那里的切线重合,因而还剩下 $6 - 2 = 4$ 条。这样就一共得到 $4 \cdot 6$ 条二重切线(因有 6 个二重点,而其中每点有 4 条切线)。当二重点转化为一般 C_5 的非奇点时,每一条这样的二重切线化为两条(像 n 为偶数时那样),于是就有 $(4 \cdot 6 \cdot 2 =)48$ 条二重切线。至于这些二重切线是实是虚,我们不了解,也不去考察。但仍然可以肯定的是：

对于近邻曲线 $\varphi\Omega = \pm\varepsilon$,这 48 条二重切线,只要它们是实的,就肯定不是孤立二重切线。因为当曲线转化回到退化曲线时,一个切点要落在二重点;另一个则落到离二重点稍远处;而对于一个孤立

二重切线，在曲线转化时则不然，那时两个切点都要落到二重点。

关于最后一种二重切线，就要考虑二重点的各个连线。这种连线共有 $\left(\frac{6\cdot 5}{2}=\right)15$ 条，但作为二重切线，每条要像偶数 n 的款那样计算 4 次。因此，在向近邻曲线转化时，6 个二重点给出 $(15\cdot 4=)$ 60 条二重切线。

我们立刻进一步指出，只要这样所得到的二重切线是实的，它们就不是孤立的，因为它们的切点总是出现在被连接的二重点附近。

这样计算出的二重切线数目是 $(12+48+60=)120$ 条。这样就对我们的曲线验证了一般理论的数字。

现在已经得到了检验近邻曲线 $\varphi\Omega=\pm\varepsilon$ 所需的情况。

关于实拐点，显然有 C_3 的 3 个拐点以及 6 个二重点中每两个所提供的，因而共有 $(3+2\cdot 6=)15$ 个。只要有实二重切线，它们都不是孤立的，因而 $t''=0$，于是在这里，$w'+2t''=15=5\cdot 3=n(n-2)$ 也正确。

任意奇数 n 的情况类似。我们还是从一个退化的 C_n 出发，它是一条 C_3 和 $\frac{n-3}{2}$ 个全等的共中心的实椭圆组成的。这些椭圆中紧挨着的两个主轴作 $\frac{2\pi}{n-3}$ 角。

对于这样得到的任意奇次 n 的无奇点曲线，总有

$$w'+2t''=n(n-2)。$$

§30.9 举例说明证明中的连续性方法，证明的完成

在作出这些例子之后，我们回到关于所要阐述的证明的一般性讨论。我们知道，在那个 $\frac{n(n+3)}{2}$ 维代表空间，从一个无奇点的 C_n

的代表点 1 到另一个无奇点的 C_n 的代表点 2 的连续变化，可能需要经过流形 $D=0$ 上的若干点。这些点代表着只具有一个二重点而无其他奇异性的 C_n，而且每个二重点把曲线的类数降低两个单位。现在我们想知道对于点 1 成立的关系 $w'+2t''=n(n-2)$ 在变动到点 2 时，有无变化；我们要考察，会不会有所减少或增加。

我们可能以为，经过具有普通二重点的曲线时，拐点数 w' 或二重切线数 t'' 会有增减。但事实不是这样，而是有定理：按所说的方式穿过流形 $D=0$ 时，数字 $w'+2t''$ 保持不变。

我们自然要考虑到：所通过的 C_n 可能有具两个实分支的二重点，也可能有孤立二重点(图 30.22)。

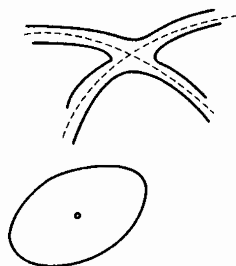


图 30.22

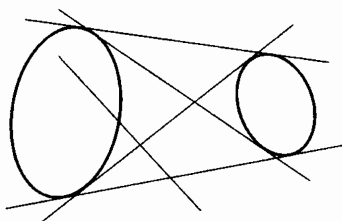


图 30.23

在前一款，在转化到具二重点曲线时，显然要失去两个实拐点，但当重新转化为无奇点曲线时，又要出现两拐点，因而 w' 保持不变。

在后一款，在转化中，根本没有拐点参与其中，因而谈不上 w' 有任何变化。

于是就解决了第一个问题：

在转化中，若经过一条具二重点的曲线，使曲线类数降低两个单位，则无奇点曲线的拐点数 w' 不变。

同样地，t'' 也不发生变化。一般地，实二重切线的数目是会变的，但数目变化的二重切线却都不是孤立的。例如考虑图 30.23，当其中一个卵形线缩成一个孤立点然后转化成虚点时，4 条二重切线

变成虚的，但前前后后都不会出现孤立实二重切线。

当一个具实分支的二重点吸收实二重切线时，情况也是如此：

在曲线的连续变化中，当越过具二重点的曲线时，实二重切线数目会有大的变化，但孤立实二重切线数目则不受影响。

把关于 w' 和 t'' 所说的归纳在一起，我们看到，在所讨论的转化中，$w+2t''$ 在转化前后是相等的。

这样就把穿过曲面 $D=0$ 的问题解决了，若穿过曲面 $D=0$ 时遇到的是代表着高级奇异性的点，事情自然要复杂得多。妙处正是在于，这种复杂性可以完全避免，若有人不嫌麻烦，肯于研究穿过曲面 $D=0$ 的高级奇异性的影响，那他的最后结论毫无疑问仍然是：$w'+2t''$ 这个数字经过高级奇异点也不变。这是因为，由始点到终点的路线总可以绕过 $D=0$ 的高级奇异点。

既然 $w'+2t''$ 在越过曲面 $D=0$ 前后保持不变。这个数目在什么情况下还有可能变化呢？对我们来说，发生变化的地方只能在 $D\gtrless 0$ 时，因为 $D=0$ 代表着比 $D\gtrless 0$ 更高层次的情况。而根据我们的一般讨论，那是可以回避的。因此，可确切地提出以下的问题：

如果所考虑的只是 C_n 的系数满足一个单一的代数条件[①]的情况（而这是合理的），又不再考虑已经解决了的特殊款 $D=0$，那么从一个代表点到另一个时，$w'+2t''$ 这个数目还有可能变化吗？

由于一条无奇点的 C_n 的拐点总数完全确定于 n，而 w' 的个数只能在无奇点的曲线上两个实拐点重合，然后变成虚拐点，或者在这个过程倒转来时，才能变化。

另一方面，由于相应原因，t'' 的个数变化也只有两种可能：或者两条孤立二重切线重合，然后变为虚的（或者倒转来）；或一条实孤立二重切线化为非孤立的（或者倒转来）。

① 原文作“代数条件方程”。——中译者

这里所说的 t''发生变化的两种可能性中的第一种是无需考虑的；因为那要涉及两个条件。事实上，若有两条孤立的二重切线重合，就或者会出现一条四重切线，或者一条切线在两个点各有二阶切触[①]（因为那两个共轭虚切点必然有相同的遭遇）。

于是 t''遇到变化只能是一条孤立二重切线化为一条非孤立二重切线；或其逆过程。

可喜的是，剩下我们还需要考虑的 w'和 t''的可能变化必然要同时出现，而且那时 $w'+2t''$ 仍保持不变。

这里要阐明的定理我们先用曲线图来说明，然后问：

我们所看到的，能变成关于理想代数曲线的严格定理吗？

我们画一段具有两个实拐点的曲线，让拐点会合然后变成虚的。在图 30.24(a)里显然有一条具两个实切点的二重切线；由此得到图 30.24(b)的一条 4 点切线；然后在 30.24(c)，从代数意义上说，有一条孤立二重切线。这是因为，一条实二重切线要变成虚的，就只能先和另一条实二重切线会合，而且目前的款却不是那样。因此：

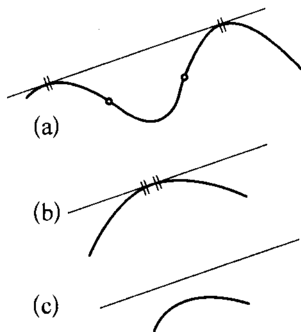


图 30.24

当两个实拐点会合而变虚时，二重切线的两个切点也要会合，然后变成虚的。二重切线本身则仍是实的。

由此可见，当 w'减小两个单位时，t''增加一个单位。但在 $w'+2t''$ 中，t''要乘以 2，故关系仍然成立。对这种情况，我还要通过以前得到过的两种形状的 C_4 来说明。

我们首先还是把以前曾讨论过的 C_4 的两种情况弄清楚。

① 或说“各有 4(虚)点的切触”。——中译者

我们曾经从两个相交的椭圆出发,并由此先作出 $\Omega_1\Omega_2=-\varepsilon$ 的曲线,它为 4 条卵线所组成(图 30.25)。现在令 ε 变大,则二重切线就要脱开,如图 30.26 所示。这时实拐点不见了,但出现 4 条实孤立二重切线,因而关系 $w'+2t''=8$ 仍然成立。

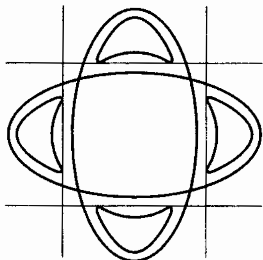


图 30.25

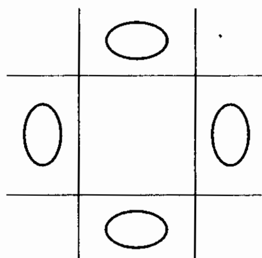


图 30.26

那以后我们令 $\Omega_1\Omega_2=+\varepsilon$ 以得到一个具有 8 个拐点的环形曲线(图 30.27),其中外面一支有 8 个实拐点和 4 条二重切线。令 ε 增加,在一定时刻,4 条二重切线就都变成 4 点切线,然后进一步变成 4 条实孤立二重切线。这时曲线的外面分支就变成处处凸的卵形线,那 4 条孤立二重切线围着它,但不和它相交(图 30.28)。同时曲线的 8 个拐点都消失了。于是在这里,关系 $w'+2t''=8$ 也是不变的。问题是,为什么是这样?为什么不只是对于 $n=4$,而是对于任意 n

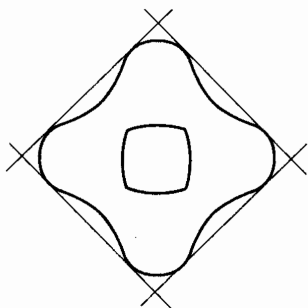


图 30.27

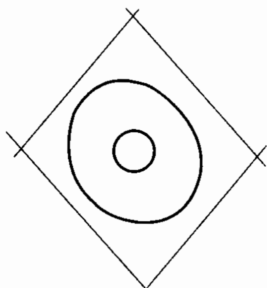


图 30.28

也是这样？对此，我不违背这个讲演的原则地说：首先是对于任意 n 所得的图和这里所讨论的是完全类似的，但需要检验，如何能从这些图形对理想代数曲线推出严格的结论。

我概略地把作法说明如下：首先考虑一些经验图形的情况。设有任何不向无穷远伸展的一段线路，它有一条二重切线，则曲线有一段在有穷远处把两个切点相连。由图 30.29 可知，在切点之间至少有两个拐点；而数目也的确可能等于 2——倒转来，若有一段本来无拐点的曲线，而把中间压挤，使它上面出现两个或更多拐点，则曲线至少有一条二重切线。

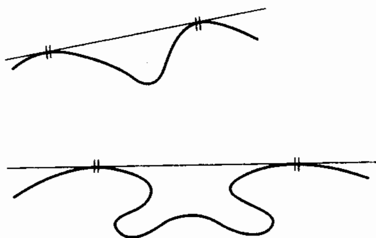


图 30.29

其次，我们要把这个结论移植到正则的理想曲线，根本方法是正确应用罗尔定理或者魏尔斯特拉斯关于连续函数的极大值存在定理。假定曲线段有一条二重切线，则选取 x 轴平行于二重切线，并把切点间的曲线段用参数方程 $x = \varphi(t)$, $y = \psi(t)$ 表示。于是在两个切点，$\dfrac{\mathrm{d}y}{\mathrm{d}x} = \dfrac{\psi'}{\varphi'}$ 等于 0，但 y 在那个节里必有极大值或极小值，而在那里，$\dfrac{\mathrm{d}y}{\mathrm{d}x} = \dfrac{\psi'}{\varphi'}$ 也等于 0。由此可知，在该节里，$\dfrac{\mathrm{d}^2 y}{\mathrm{d}x^2}$ 至少两次等于 0，而这就是所要证明的，类似地可以得到逆定理。

第三，我们认为，对于代数曲线，若把它们限在“一般”的情况内，则当在一条二重切线两个切点之间的曲线弧逐渐变形，使两切点重合时，这段弧只有两个拐点。因为若那里有多于两个的拐点，则令曲线段变形，使两个切点重合时，就得到高级的情况，而那是我们一般讨论所排除的。

把这 3 点归纳起来，就可以推知：对于理想代数曲线，经过所讨

论的变形，$w'+2t''$ 是常数。因此，对于无高级奇异性的曲线 C_n，$w'+2t''$ 有常数值 $n(n-2)$。证毕①。

这里所阐述的定理证明和 *Math. Ann.* 第 10 卷中所给出的证明，其区别只在于：那里更多地直接利用了经验曲线，这里则更深入地说明了由图形到理想代数曲线的移植。

① [在第 208 页所举的克莱因论文中还证明了

$$n+w'+2t''=k+r'+2d''。$$

其中，n, w', t'', k, d''的意义我们都已了解，r'是实尖点数。与此有关，布里尔(*Math. Ann.* 第 16 卷，1879 年，第 348—408 页)指出了高级奇异性出现的情况；和上述相反，他的研究是纯代数的。可以参考于尔(*Math. Ann.* 第 61 卷，1905 年和 *Bericht über den sechsten Skandinavischen Mathematikerkongress in Kopenhagen*，1925 年，第 119—126 页)。舒(Fr. Schuh)在下列文章中，把问题推广到复曲线：*On an expression for the class of an algebraic plane curve with higher singularities* (《具高级奇异性的平面代数曲线的类数表达式》，Akad. d. Wiss. Amsterdam，1904 年，第 42—45 页)。若 $\sum u$ 是具实点的一切奇异性之和，$\sum v$ 是具实切线的一切奇异性的类数之和，舒得到

$$n+\sum v=k+\sum u。$$

继续深入克莱因的研究的最新工作是霍尔克罗夫特(T. R. Hollcroft)的"Reality of singularities of plane curves"("平面曲线的实奇异性"，*Math. Ann.* 第 97 卷，1927 年，第 775—787 页)，那里讨论了一条具有已给次的代数曲线的实尖点的最大数问题。]

第九部分

用作图和模型表现理想图形

§1 无奇点挠曲线，特殊地，C_3 的形状（曲线的投影及其切线曲面的平面截线）

这个讲演里前面所持的一个主要观点是，把从经验得来的空间直觉及其有限度的准确性和精确几何的理想化表述加以区别。一旦了解了两者的区别，我们就可以专门选择这条或那条途径。一种可能性是放弃严密的概念构造，专以经验的空间直觉所得事实为基础建立一种几何，这时谈的就不再是点和线而总是“斑”和“带”。另一种可能性是，把空间直觉看做不可靠而完全放在一边，只考虑用纯解析方法所获得的抽象结论。实际上，这两种办法都不能取得丰硕成果；我总是主张，在弄清楚两种方向的区别之后，把它们结合起来。

把两者结合会产生惊人的力量。因此，我一贯强调，抽象情况也要利用实物模型来显示，而这种想法也正是产生我们哥廷根那套模型的根源。我想用这个讲演的最后几小时来介绍几种有趣的模型并作相应的解说。

我立刻从空间几何开始，先谈挠曲线。

在精确几何里，挠曲线是由 3 个方程

$$x=\varphi(t),\ y=\psi(t),\ z=\chi(t)$$

确定的正则理想曲线，其中 t 在一个节里变动，假定 $\varphi,\ \psi,\ \chi$ 是不全等于常数的连续函数，两次可微而且其二阶导数是分段单调的。

这种曲线有“切线”和“密切面”。

我们知道，在点 $x_0=\varphi(t_0),\ y_0=\psi(t_0),\ z_0=\chi(t_0)$ 的切线方程可以写成

$$\begin{Vmatrix} x-\varphi(t_0) & y-\psi(t_0) & z-\chi(t_0) \\ \varphi'(t_0) & \psi'(t_0) & \chi'(t_0) \end{Vmatrix}=0。$$

令“矩阵”等于 0 自然就是表示下面方程组成立

$$\frac{x-\varphi(t_0)}{\varphi'(t_0)}=\frac{y-\psi(t_0)}{\psi'(t_0)}=\frac{z-\chi(t_0)}{\chi'(t_0)}。$$

在这一点的密切面方程是

$$\begin{vmatrix} x-\varphi(t_0) & y-\psi(t_0) & z-\chi(t_0) \\ \varphi'(t_0) & \psi'(t_0) & \chi'(t_0) \\ \varphi''(t_0) & \psi''(t_0) & \chi''(t_0) \end{vmatrix}=0。$$

此外,有以下定理:

若作一条弦连接挠曲线的两个点,并令两点以任意方式重合,则弦的极限位置是切线。

若作平面经过挠曲线的 3 点,并令 3 点以任意方式重合,则平面的极限位置是密切面。

现在问题是,这些定理在模型上显示出的形象是如何的呢?

我从向你们提供几个挠曲线模型开始。这里当然有这样一个特殊困难:我只能限于描述现象,而要求你们随后自己对模型观察所谈论的事实。我的讲解也只好是说教式的。

这里是一条一般无奇点的挠曲线模型。在曲线"充分小"的一段 k 上取任意一点 p。我们的主要问题是:从任意一点 O 把 k 投射到任意不经过 O 的平面上,k 的投射 k' 的形状如何?我们的答案,按照 O 的不同位置而有 3 种:O 不在密切面上,在密切面上但不在点 p 的切线上,以及在 p 的切线上。设 p' 是 p 的投影,则有:

(1) 当 O 不在 p 的密切面上时,k' 是一段曲线,它在 p' 没有奇异性(图 1)。

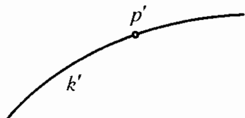


图 1

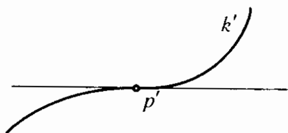


图 2

(2) 当 O 在 p 的密切面上时，由于挠曲线穿过密切面，投影 k' 在 p' 有拐点(图 2)。

(3) 当 O 在 p 的切线上但不和 p 重合时，k' 在 p' 有尖点(图 3)。这通过下述方法即可看出：

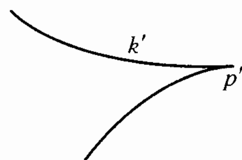


图 3

先选取 O 不在 p 的切线上而在一条弦 pp_1 上，再令 p_1 无限制地趋于 p。这时所得投影是一条有二重点的曲线。在二重点的两条切线，按照 p_1 的选择，对应于曲线在 p 和 p_1 的两条相错切线。在完成极限 $p_1 \to p$ 时，在二重点的两条切线重合为一，而二重点本身就趋于尖点。

现在，结合代数三次挠曲线，关于挠曲线的这种情况就可以得到基本全面的了解。因为作为代数曲线，可以对它的投影的整体而不限于一小段来考察。

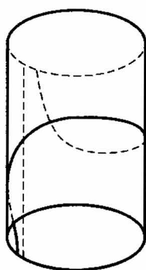
在图 4 里，画的是一条所谓的椭圆挠线，它可以看做一个椭圆柱面同一个与它有一条公共母线的斜圆锥的交线；或者更确切地，这个交线的一部分。容易看出，这样的曲线必然是三次的，即一个平面必然和它有 3 个交点，其中两个可能是共轭虚点。因为根据贝祖定理，两个二次曲面交于一个四次曲线。现在它们有一条公共的直母线，这四次曲线就要退化为这条母线和一条三次挠曲线。图 5 表示这条三次曲线以及它的切线所构成的可展曲面①。

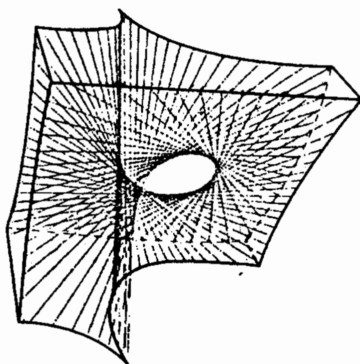
为了更好地理解下面所要讲的，先不加证明地叙述一个定理：经过空间每点，可以作一条直线和我们的 C_3 交于两点。这两个交点可以：

---

① [图 5 中椭圆挠线的切线曲面的线条模型是路德维格(W. Ludwig)制作的，由莱比锡的 M·席林(Verlage von M. Schilling) (M·席林出版社)展出。参考模型附件：路德维格。Die Horopterkurve mit einer Einleitung in die Theorie der kubischen Raumkurve (《霍罗普特尔曲线②，附三次挠曲线引论》)。也由 M·席林莱比锡展出。]

② 霍罗普特尔是 Horopter 的音译。两个射影相关的直线的对应直线交点的轨迹是一般三次挠曲线；若射影关系是全等关系，则所得的是霍罗普特尔曲线。——中译者


图 4


图 5

(1) 实而不同;

(2) 实而重合;

(3) 共轭虚而不同。

在第一款,我们有一条普通的弦;在第二款,有一条切线;在第三款,有一条所谓的"伪"弦,但作为共轭虚点的连线,却是实的。现在,考察 C_3 的切线曲面,就可以看到,它把空间分成两个域:一个域夹在切线曲面两叶之间,从其中的点作的弦是"伪"的;从另一个域的点作的弦是普通的;经过切线曲面上的点,则有曲线的切线。现在,以切线曲面外的任意点为中心,对 C_3 作投射,则投射直线构成具有一条二重棱的三次锥面。投射锥面是三次的,因为每个平面和 C_3 交于 3 点,因而经过投射中心的每个平面和投射锥面交于 3 条直线。那二重棱就是从投射中心所作的 C_3 的弦,因而总是实的,但对于两个区域中之一,它是孤立弦。现在,若令一个平面和投射锥面相交,则当锥面有一个普通二重棱时,交线是一条具有一个普通二重点以及一个拐点的平面三次曲线。C_3 上和这个拐点对应的点是 C_3 上具以下性质的点,即在那个点,C_3 的密切面经过投射中心(图 6(a))。但若投射中心在有"伪"弦的那个域内,则投影曲线是具有孤立二重点和 3 个实拐点的平面三次曲线。因

为这时在 C_3 上有 3 个点，在那里的密切面都经过投射中心（图 6(b)）。由图 6(a)到图 6(b)的过渡中，若投射中心落在切线曲面上但不在 C_3 上，则投射锥面在二重棱的两个切面合成一个二重面；投影曲线有一个尖点，仍保留一个拐点。尖点所对应的是 C_3 上的一个点，那里的切线经过投射中心（图 6(c)）。

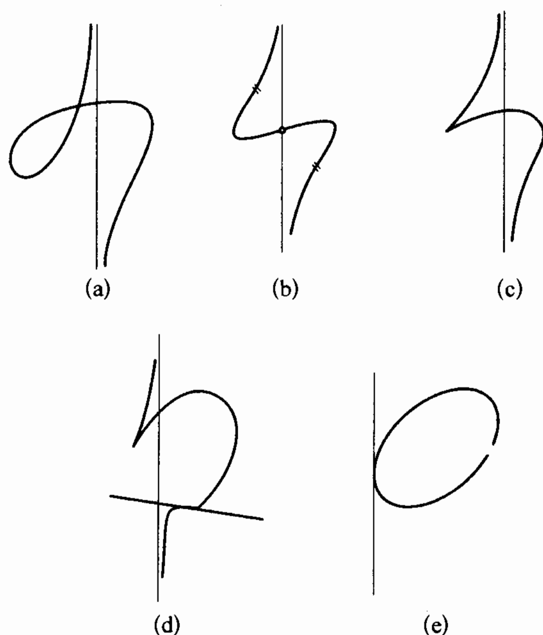


图 6

到此为止，困难不大；因为这 3 个图显然是彼此连续地联系着的。比较困难的是下面的问题：

当投射中心在曲线上时，投射曲线作何形状？

这时，从代数观点来看，显然投影曲线必然要退化为一个二次曲线和一条直线；投射三次锥面分解为一个二次锥面和投射中心所在的密切面。但由连续变化观点，投影曲线中的这条直线是怎样连续地从曲线中分离出来的呢？我说，图 6(c)转化为图 6(d)，然后进一步转

化为图 6(e)。显然这 3 个图也可能用和它们射影等价的图代替[①]。

这些已经是很值得注意的图形。更值得注意的是挠曲线 C_3 的切线曲面的平面截线。这个可展曲面是挠曲线的密切面的包络，即密切面沿着它所含的切线和切线曲面相切。挠曲线构成切线曲面的“脊线”。沿着挠曲线，切线曲面就“像剃刀那样锋利”。

无论如何，像模型所示那样，一条接一条切线会构成一个曲面的刀锋，乍一看是会令人惊异的。

一般来说，考虑切线曲面的平面截线，和考虑挠曲线的投射锥面是对偶的——每一个了解射影几何的人都会这样说。事实上——我不深入说明，一般地有以下对偶[②]：

可展曲面——挠曲线；

可展曲面的平面截线——挠曲线的投射锥面；

可展曲面的母线——挠曲线的切线。

特殊地，对于我们的三次挠曲线，曲线上的点和它的密切面所产生的图形也是对偶的，但我们不作进一步的讨论。

因此，要问切线曲面的平面截线，我们只需取上面所作的图 6(a)—图 6(e)的对偶。

图 6(a)中曲线的类数是 4，因为一般三次曲线的类数是 6。而曲线的二重点把类数降低两个单位，其次数自然是 3，于是，我们有结论：

切线曲面的任意平面截线是四次三类曲线，它有一个二重切线，对应于投影曲线的二重点(图 7)。

我们知道，一条二重切线可以有实切点或虚切点。因此，截线有两种，或者像图 7(a)那样，有一条具实切点的二重切线，这时曲线有一个尖点，或者像图 7(b)那样，有一条具虚切点的二重切线，这时实

① 图 6(d)中的拐点(在拐点的切线是经过投射中心的密切面的切点的投影)不可能和尖点重合，否则投影曲线就会和一条直线(在拐点的切线)交于 5 点，拐点计算 3 次，尖点 2 次。

② [参看第二卷第 63—65 页(中译本第 68—70 页)。]

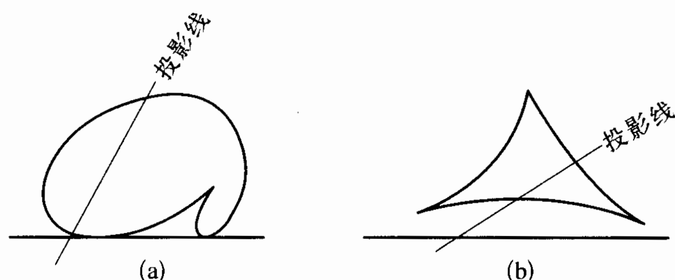


图 7

曲线有 3 个尖点。这两图是上面图 6(a)和图 6(b)的对偶。

若带有这样的了解来考察那个模型，我们先是看不到切线曲面有哪种形状的截线。我们还必须考虑所画曲线经过射影变换的形状。这情况就像 3 种看起来形状不同的二次曲线（椭圆、双曲线、抛物线）在射影几何观点看来是等价的。

为了获得“双曲”型的截线，我们在新作的图 7(a)和图 7(b)里画了一条和曲线相交的直线，并取它作为适当射影变换的投影线。这样就得到两个曲线。它们各有两条渐近线，曲线的分支沿着它们的方向延伸到无穷远处①。

如果定性地作图，就得到图 8(a)和图 8(b)那样形状的两条曲线。它们实际上代表切线曲面模型在适当平面上的截线，那个模型有延伸到无穷远处的两叶（见图 5）。

另一个问题是，当截平面经过挠曲线的一条切线时，截线的形状如何？

在此情况下，显然切线曲面的这条母线要从截出的四次曲线分离出来，剩下还有一条三次三类曲线，即具一个尖点的曲线（图 8(c)）。进一步考察表明，它和我们的切线相切，而切点也就是该切线

① [图 7(a)，图 7(b)本身（不考虑所画的截线）表示椭圆形截线（投影线和曲线不相交也不相切）。]

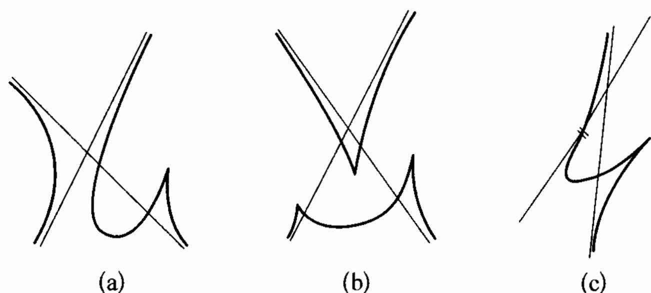


图 8

和挠曲线的切点。

这些结论是容易抽象地得到的。但还可以问：图 8(a)是如何通过图 8(c)连续地过渡到图 8(b)的？这对应于由一条具二重点的三次曲线到具尖点的三次曲线的过渡，请你们在模型上验证我下面的说明：

我们由图 8(a)或图 8(b)得到图 8(c)如下：像图 9(c)所示那样，令图 8(a)中的两个分支 1 和 2 逐步接在一起；或像图 9(b)那样，令图 8(b)中(参看图 9(b))属于不同分支的两个尖点相互接近。也可以倒过来由图 9(c)得到图 9(a)和图 9(b)，只要令图 9(b)中的拐点切线以这样或那样的方式化为曲线的一部分。

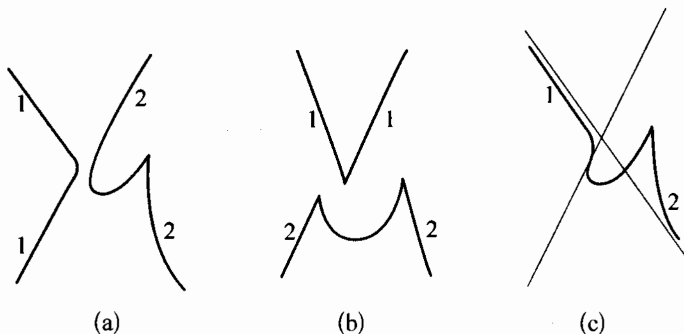


图 9

最后考虑切线曲面和一个密切面的交线，我们得到一个二次曲线和一条计算两次的切线（图 10）。这个图是由图 9 经过图 11 的过渡位置转化而来，显然，令图 9(c)中的拐点切线固定，把曲线上的分支 1 向无穷远伸展的部分向下弯，同时将分支 2 的那部分向上转，使它的尖点指向下方，就得图 11，再令尖点和拐点相合，就得图 10。

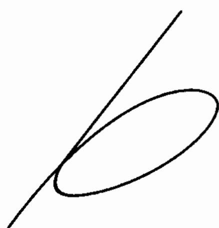


图 10

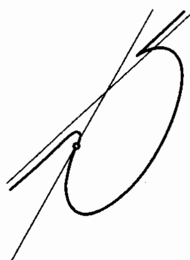


图 11

还可以指出，这样的过渡在第 242—243 页已谈论过了。

在说明这些之后，还要请你们每位通过模型验证一下，使直观形象更加生动。

这里谈到的空间曲线，特别是三次挠曲线的射影性质，当然是和数量的考虑相联系着的。在这里，我只好请诸位参考关于曲率和挠率以及曲线展开的理论。关于最后一点，我只指出，当我们让一个平面绕一个可展曲面展开，或一条直线绕一条曲线展开时，就得到密切面渐伸线[1]和切线渐伸线（即一般渐伸线）[2]。

---

① Planevolvente 的暂译。当一点沿挠曲线运动时，曲线在该点的密切面上一个定点的轨迹叫做曲线的密切面渐伸线。与之相应，切线上一个定点的轨迹就叫做切线渐伸线(Filarevolvente)（简称渐伸线）。——中译者

② 对这些内容作出特别全面讨论的教材可以指出：谢弗思：*Einführung in die Theorie der Kurven in der Ebene und in Raume*（《平面和空间曲线引论》），莱比锡，1900 年，第 3 版，1923 年。

[Lilienthal, R.：*Vorlesungen über Differentialgeometrie* Ⅰ（利林塔尔：《微分几何讲义》Ⅰ）莱比锡，1908 年。]

§2 挠曲线的7种奇点

我进一步谈谈挠曲线的奇点。在平面上,可以列一个表来表明:沿曲线前进时,曲线上的点(p)和切线(t)向前或向后[①]运动。向前时用符号(+)表示,向后用(−)表示。这样得到含4种情况的表(见表1),代表着普通点以及3种不同的奇点(图12)。

表 1

p	t	
+	+	普通点(1)
+	−	拐点(2)
−	+	普通尖点(3)
−	−	鸟喙尖点(4)

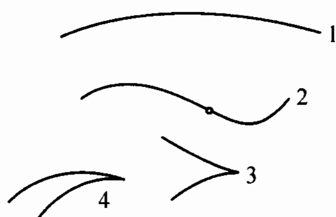


图 12

1847年v·施陶特在他的*Geometrie der Lage*(《形位几何》)(第113页)里对三维空间绘出了相应的表。这时除曲线上的点(p)和切线(t)外,还要考虑密切面(e),于是可以用符号(+)或(−)作成含8种组合的表(见表2)。

表 2

p	t	e
+	+	+
+	+	−
+	−	+
−	+	+
+	−	−
−	+	−
−	−	+
−	−	+

① "向后"即和原来方向相反。——中译者

这样就对挠曲线纯形式上得到 7 种不同的奇点，但它们并没有特殊的名称，不过它们的特征是清楚的。对这 8 种情况，维纳都作了模型①。

为了使我们的说法易于理解，我暂时回到平面曲线。取曲线上考虑中的点为一个直角坐标系的原点，切线为 x 轴，则上面所说平面曲线的 4 种情况，显然分别是，若曲线从第 1 象限开始继续行进，就进入第 2，第 3，第 4 或回到第 1 象限，如图 13 所示。

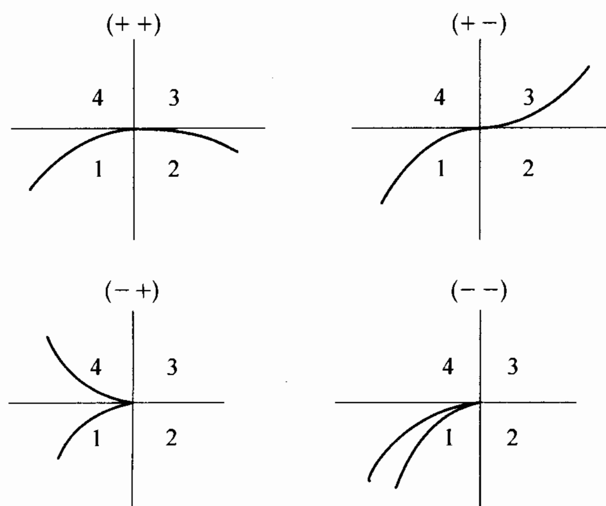


图 13

可以设想，挠曲线的 8 种情况也可以通过曲线在坐标系 8 个卦限中的不同走向来描述。事实上也的确如此。具体地可以这样做：取直角坐标系，以考虑中的点为原点，以那里的切线为一条坐标轴，以那里的密切面为一个坐标平面。这样，若曲线从第一卦限开始，则它继续行进时，恰好有 8 种可能性，即看它进入 8 个卦限中的哪一个，而曲线的这 8 种不同形状也正是上文根据 v · 施陶特的结果所列出的。

① ［参看 *Zeitschr. f. Math. u. Phys.* 第 25 卷，1880 年，第 95—97 页。这些模型可以从 M · 席林莱比锡得到。］

这8种形状具体如何，需要利用模型来考察；因为在表中，它们是平行列出来的，对曲线形状没有说明。此外，在模型中还可以从不同观察点以得到曲线的投影；上面已经谈到了最简单的情况（+++）的投影。这个观察点（即投射中心）可以：不在密切面上，或在密切面上，或在切线上。这里值得注意的是，对于一定的观察点，挠曲线上的强奇点在投影上可能消失。例如在（－++）的款，若观察点在切线上，其投影是（++）型的平面曲线，它没有奇点，或者说，有一个"掩盖了"的奇点。值得注意的另一个结果是：对于曲线上（－－－）的点，无论观察点在哪里，所得投影曲线总有一个鸟喙状尖点。

此外，我还想鼓励人们把关于曲线形状的这种考察进一步深入下去。特殊地，可以指出这样的课题：考察挠曲线的切线曲面在8个款中的形状，特别是它的平面截线是什么样的？它们之间又是如何连续地互相联系着的？

关于这里所论的内容，我略谈解析方面的文献，这些作者采取的方法是把 x，y，z 展开为 t 的幂级数。这方面著作有[①]：

H. B. Fine：*Über Singularitäten von Raumkurven*（芬恩：《关于挠曲线的奇点》）。Dissertation Leipzig，1886；以及 *American Journ. of Math.* 第8卷，1886年，第156—177页；

O. Staude：*Über den Sinn der Windung in singulären Punkten von Raumkurven*（施陶特：《关于挠曲线在奇点挠率的意义》）。同上，第17卷，1895年，第359—380页；

A. Meder：*Über einige Arten singulärer Punkte von Raumkurven*（梅德尔：《关于挠曲线的几种奇点》）。*J. f. Math.* 第116卷，1896年，第50—84，247—264页。

应当把我希望看到的关于图像形状的讨论和这些著作相结合。

---

① [此外，可以参考第247页所列的利林塔尔的书第255—272页以及梅德尔"Analytische Untersuchung singulärer Punkte von Raumkurven"（"挠曲线奇点的解析研究"。*J. f. Math.* 第137卷，1910年，第83—144页。）]

这样就结束了我对挠曲线的阐述。下面我转而讨论曲面。

§3　关于无奇点曲面形状的一般讨论

若回忆一下我们对平面曲线已作出的种种区别(解析的、非解析的、正则的,等等),就完全有理由预料,在这里只能涉及特款。为简明起见,我限于谈论代数曲面。

一个代数曲面用方程

$$f(x, y, z) = 0$$

确定,其中 f 是多项式。

我们习惯把 f 中同次的项放在一起,于是曲面方程又可以写作

$$0 = f_1 + f_2 + \cdots$$

为了考察曲面在一点邻近的情况,我们利于取该点为坐标原点,这样,常数项 f_0 就不出现。我们看到,一般地,可以把 $f_1 = 0$, 即曲面在点 O 的切面,看做曲面的一阶近似。我们还可以把 $f_1 = 0$ 作为切面的定义:令一次项等于零,就得到切面方程。

若要得到较高阶的近似,就令

$$0 = f_1 + f_2,$$

这个曲面就叫做密切二次曲面,如此等等。

在这里我们立即得到曲面上点的分类:若 f_1 不恒等于 0,就有一个确定的切面存在,那个点就是曲面的"简单"点或普通点;若 $f_1 = 0$ 而 $f_2 \neq 0$①, 我们就有一个二重点;一般地,若 $f_1 = f_2 = \cdots = f_{v-1} = 0$ 而 $f_v \neq 0$, 我们就有一个 v 重点。下面又把二重点叫做结点。

---

① 即 f_1 恒等于零而 f_2 不恒等于零,下仿此。——中译者

现在,对于普通点,可以说些什么呢?

首先我们问,在普通点,切面和曲面交于什么样的曲线?这条截线的一阶近似是 $f_1=0$, $f_2=0$, 而这是切面和密切二次曲面的交线。由于 $f_2=0$ 代表二次锥面, $f_1=0$ 代表平面,它们的交线是一对实的或虚的,或重合的直线。根据直线 $f_1=0$, $f_2=0$ 的不同情况,可得曲面和切面交线的不同情况(图 14、图 15):

(1) 直线是实的:交线有具实分支的二重点;

(2) 直线是虚的:交线有具虚分支的二重点;

(3) 直线重合:交线有尖点。



图 14

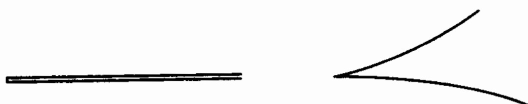


图 15

重复一下:除 $f_2=0$ 的高阶情况外,切面和曲面交于一条曲线,或者有二重点,其分支或是实的或是虚的,或者有尖点。这 3 种情况,依次称为曲面有双曲形、椭圆形或抛物形曲面曲率。

此外,不难找到不同的优美例子来显示切面和曲面的交线有更高阶的奇点,特别是可以探究诸如什么时候交线上出现鸟喙状尖点的问题。特殊地,若 $f_2=0$, 则切面和曲面有三重或更高阶奇点。与此有关的是下面问题。

我们知道,在初等微积分教材中,讨论了二元函数的极大值和极小值问题。这涉及对一个曲面 $z=f(x, y)$ 和它的一个水平切面的交线的考察。

现在，假定交线有一个鸟喙状尖点[①](图 16)，并假定图中有阴影的区域表示那里曲面在切面之上，而其余区域则表示曲面在切面之下，则显然曲面所代表的函数 $z=f(x,y)$ 在切点 O 既没有极大值，也没有极小值。

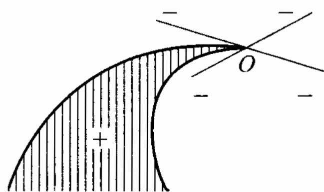


图 16

但是，若考虑经过 O 而垂直于切面的一切平面和曲面的交线，则在每个方向的平面上总有一个真正的极大值，因为截线离开点 O 时，开始总是下降的。这是由于，在 O 的紧邻，截线并不立刻就进入有阴影的区域。由此可见，对于二元函数 $z=f(x, y)$，只考虑上述的曲面的平面截线，还不能判断它是否有极大或极小值。只有当曲面与每一条经过 O 的曲线在 O 有极高点或极低点，函数在 O 才有极值。

可是当这个问题用解析方法处理时，就显得难懂。而且事实上，在早先的教材里都有着错误的结论，直到 G · 皮亚诺才把事情弄清楚，引起人们的注意。此外，魏尔斯特拉斯在他的变分法的讲演中，也介绍了正确的理论。

我们当然也可以进一步考虑曲面的数量关系，对于 $f_1 \neq 0$ 和 $f_2 \neq 0$ 的点，L · 欧拉所建立的曲率理论占有特殊的位置。

在所论的点作曲面法线，然后取经过法线的平面和曲面的截线的密切圆。这样就有两个互相垂直的法截线，它们的曲率有极大极小值。对此我不多谈，但可以指出，为了显示在曲面一个普通点的曲率理论，维纳[②]也制作了模型，绘出了各个法截面的密切圆。

---

① Schnabelspitze. ——中译者

② ［参看 W. Dyck: *Spezialkatalog der Math. Ausstellung*(迪克:《数学展览会的特殊目录》)(Deutsche Unterrichtsausstellung in Chikago, 1893, 德国数学展览, 芝加哥), 柏林, 1893 年, 第 52 页。］

最后我认为迪克的目录[①]中称为“豆状试验体”的模型值得注意，所要求的是在它上面把具双曲曲率、椭圆曲率和抛物曲率的点分别标出。对此，我们的目测显得很不确切，初等曲率理论中的定义已显出非常抽象的性质。当我们对人头雕塑像那样复杂的、由经验确定的曲面，试图按上述要求去做时，就会对此有特别生动的感受。我欢迎人们对此加以研究。

§4 关于 F_3 的二重点，特别是它的二切面重点和单切面重点

现在我谈论一个曲面上的奇点，特别是三次曲面上的，按照上面所说，若一个 F_3 在坐标原点有二重点，它的方程就可以写成

$$0 = f_2 + f_3。$$

我们假定 $f_2 \neq 0$，这样曲面就的确有一个二重点，而不是例如三重点。

我们通常说，二次锥面 $f_2 = 0$ 代表着曲面在点 O 的“一阶近似”。这个锥面可能是：

(a) 一个(实或虚的)正常锥面；

(b) 分解为两个(实或虚的)平面的锥面；

(c) 退化为一个二重平面的锥面。

与此相应，曲面有：

(a) 一个普通二重点，它或者是一个实锥面的顶点，或者是孤立点；

(b) 一个二切面重点；

(c) 一个单切面重点。

在(b)款中，两个平面的交线总是实的，叫做二切面点重的轴。

① [同 253 页注②，第 54 页。]

问题是：具不同类型二重点的三次曲面的形状是怎样的？

我首先指出：求曲面 $f_2 + f_3 = 0$ 和 $f_2 = 0$ 的交线等于求 $f_2 = 0$ 和 $f_3 = 0$ 的交线，而 $f_3 = 0$ 代表一个三次锥面。根据贝祖定理，这两个锥面有($2 \cdot 3 =$)6 条共同母线：经过三次曲面一个二重点，曲面上有 6 条直线。当然要注意，这些直线可以是实的或虚的，还可以重合。

现在，我给出一些 F_3 的例，它们分别有各种类型的二重点，但不加证明。在一般款中情况也类似。

为了得到一个具普通二重点的曲面，我们限于回转曲面。作为曲面在 xz 平面上的经线，取一条对 z 轴对称，具一个二重点，而以 x 轴为所谓拐渐近线的三次曲线(图 17)。这样，曲面的形状就完全清楚了。但经过二重点的 6 条直线显然都是虚的。完全类似地可以作一个具孤立二重点的曲面。我们只需画一个具孤立二重点的三次曲线作为经线(图 18)。

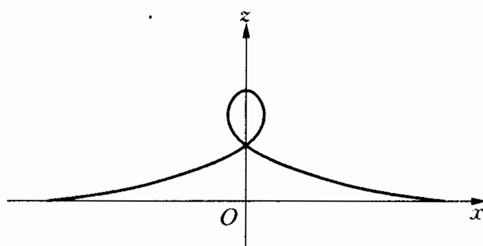


图 17

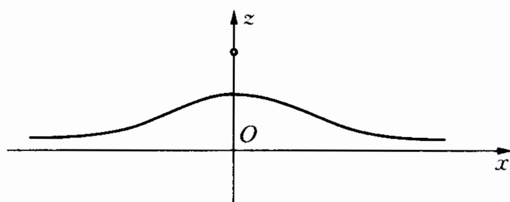


图 18

其次，我考虑这样一个曲面，在它上面经过一个二重点的 6 条直线是实的。我取的是一个具有 4 个普通二重点的曲面。为了便于得到它的解析表示，取这些二重点为坐标四面形的顶点。若适当地选取确定四面形坐标中的有关常数，我们的曲面方程就可以写成

$$\sum_i \frac{1}{x_i} = \frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \frac{1}{x_4} = 0,$$

乘以 $x_1x_2x_3x_4$，就得含 x_1，x_2，x_3，x_4 的三次方程

$$x_2x_3x_4 + x_1x_3x_4 + x_1x_2x_4 + x_1x_2x_3 = 0,$$

最后利用公式

$$x_1 = x,\ x_2 = y,\ x_3 = z,\ x_4 = 1,$$

通过对曲面的射影变换把四面形坐标转化为普通坐标，就得曲面方程

$$yz + xz + xy + xyz = 0。$$

这时有一个二重点在 O，其余 3 个在无穷远处。

从这样写出的方程立刻可以得到以坐标原点为顶点的两个锥面 $f_2 = 0$ 和 $f_3 = 0$。在 O 邻近，和曲面有一级近似的锥面

$$yz + xz + xy = 0,$$

它显然含 3 条坐标轴在内。三次锥面

$$f_3 = xyz = 0$$

为 3 个坐标平面构成，即 $f_3 = 0$ 和 $f_2 = 0$ 的 6 条公共母线是 3 条坐标轴，每条计算两次。于是得下述定理：

我们曲面上经过结点[①]的 6 条直线分成 3 对，它们分别同坐标轴重合。

你们非常容易证明。

---

① Knofenpunkt，即普通二重点。——中译者

沿着曲面上每条直线，曲面的切面和锥面 $yz+xz+xy=0$ 的切面相同，因而是固定的。

由于这些定理的射影性质，它们也适用于具 4 个有穷结点的曲面 F_3。这可以从模型验证（图 19）。首先显示：二重点所构成的四面形的 6 条棱都整条在曲面上，而曲面沿它们的切面都是固定的；此外，从某个水平面和曲面的交线中又分出 3 条直线（在图中只能看见其中一条），可是曲面沿 3 条直线的切面却绕它们转动。

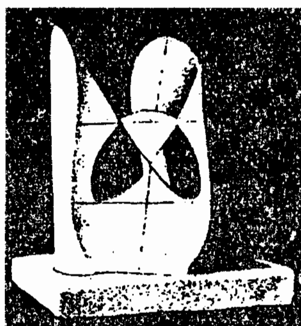


图 19

现在我转而谈 F_3 的二切面重点。

讨论这个问题已不像上面那么简单，我们对二切面重点形状还缺乏直接地了解，只能具体地加以探究。

在方程

$$f_2+f_3=0$$

中，设 f_2 是两个不同的一次式因子之积。令 $f_2=0$，就得到两个不同的平面。先假定两个因子是共轭虚的，例如

$$f_2=(x+\mathrm{i}y)(x-\mathrm{i}y),$$

则 z 轴是这两个平面的实交线，因而就是二切面重点的轴。再设曲面本身方程是简单的

$$0=x^2+y^2+z^3。$$

这时二切面重点是什么样的？如果说，曲面和两个共轭虚平面一阶近似，在这一定意义上是不错的，但不好理解。我通过下面的方法来说明：

我们的曲面是回转曲面。先取它在 xz 平面上的经线

$$0=x^2+z^3。$$

这是一条三次曲线，尖点在 O(图 20)。把它绕 z 轴旋转，就得到曲面[①]。一般地，有：

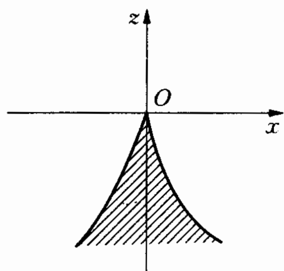


图 20

若一个二切面重点有两个虚平面，则曲面在那里的形状像一个尖锐的刺。

具有实平面的二切面重点较难说明。

我们取

$$0 = x^2 - y^2 + z^2$$

为例，它的二切面重点的平面是

$$x + y = 0,\ x - y = 0,$$

即经过 z 轴的两个互相垂直的平面，它们和 xz 平面分别成 45°角和 135°角。z 轴仍是二切面重点的轴。

令这平面之一和曲面相交，就得

$$f_3 = 0,\ 即\ z^3 = 0。$$

因此，这两个经过曲面二切面重点的纵向轴的平面都和曲面交于计算 3 次的水平直线。

若 $f_3 = a_1 z^3 + a_2 z^2 x + \cdots$，则令 $y = \pm x$，从曲面方程得到一个含 z/x 的三次方程，因此：二切面重点的两个平面中的每一个和曲面交于 3 条直线，其中两条可能是共轭虚的。

经过 z 轴的平面的截线很值得注意。当 $f_3 = z^3$ 时，平面 $y = 0$ 上的截线是 $x^2 + z^3 = 0$，即上面所得的具有一个向上尖点的曲线。平面 $x = 0$ 上的截线则是 $-y^2 + z^3 = 0$，即一条有向下尖点的平面曲线(图 21)。因此，若令平面 $y = 0$ 转到 $x = 0$，则尖点从向上变成向下；其过渡是在平面 $x + y = 0$ 和 $x - y = 0$ 上，而这两个平面上的截

① 所以这个曲面是由第 255 页两图所产生的曲面的过渡。

线已经讨论过了。于是有定理：

经过二切面重点的轴的平面一般地和 F_3 交于一个具有尖点的曲线，而尖点切线是 z 轴，只有当该平面是二切面重点的切面时，其交线才是 3 条直线。

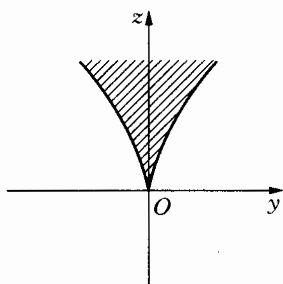


图 21

二切面重点的平面把经过轴的平面束分为两半，其中一半的平面和曲面的截线尖点向上，另一半这样的尖点向下。

图 22 显示这样一个二切面重点：它的每个切面和它相交于 3 条实直线，在这两个切面的邻近，经过轴的平面和曲面的截线，除有一个尖点外，还有两个分支。

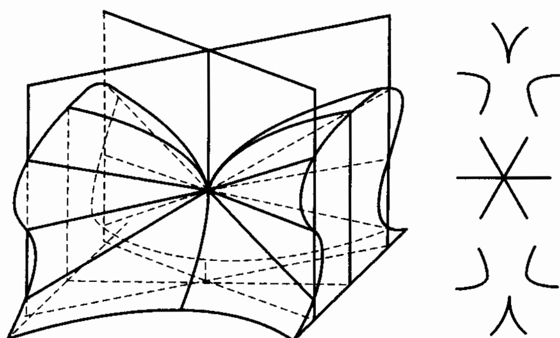


图 22

上面谈到的二切面重点的形状，不经过仔细考察是不容易说清楚的，单切面重点也是这样①。

这时，“作为一阶近似”的是一个二重面，它用二次项代表，我们

① 关于二切面重点和单切面重点的形状大约最早为库默尔和施莱夫利(Schläfli)所给出，后来我在论文“Über Flächen dritter Ordnung”(“关于三次曲面”)中用到了(*Math. Annalen* 第 6 卷，1873 年，第 551—581 页)[加上 F·克莱因和弗尔迈尔斯(H. Vermeils)的补充后，重印在 F·克莱因：*Gesammelte Math. Abhand-lungen*(《数学论文集》，第 2 卷，第 11—62 页。)]

简单地令它为 $z^2=0$。于是曲面方程是

$$z^2+f_3=0;$$

为简单起见,我们选取 f_3,使它只含 x, y,并分两款(其意义很快即可明了),第一款是

$$f_3=x(x^2-y^2),$$

第二款是

$$f_3=x(x^2+y^2)。$$

在第一款,锥面 $f_3=0$ 和 $f_2=0$ 交于 3 条实直线,在第二款,交于一条实的和两条虚直线(图 23)。

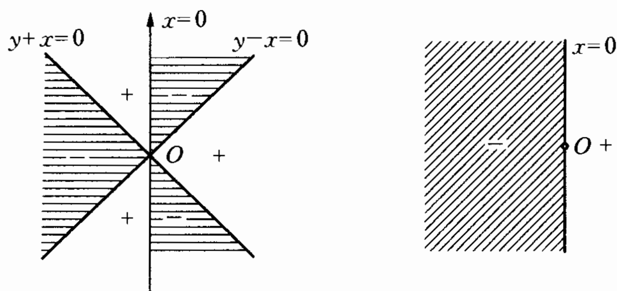


图 23

一般地有定理:

单切面重点的平面和 F_3 交于经过 O 的 3 条直线,其中 1 条或者 3 条都是实的。

关于 z,由 F_3 这方面可知,在第一款,有

$$z=\pm\sqrt{-x(x^2-y^2)},$$

在第二款,有

$$z=\pm\sqrt{-x(x^2+y^2)}。$$

除 $f_3=0$ 外,只有当 f_3 有负值时 z 才是实的,而且在 f_3 有

负值的域里，对于每对 x，y 值，z 有两个实值。因此，从上方向下看，曲面覆盖图 23 中 xy 平面有阴影部分两次，而根本不覆盖其余部分。

对这些我还要加以下几句说明：

若单切面重点的计算二次的平面和 F_3 交于 3 条实直线，则 F_3 沿这 3 条直线穿过切面，因而在 xy 平面上 3 个角度里，从上方向下看，曲面覆盖 xy 平面两次，经过 O 而不和 xy 平面重合的平面上的截线在 O 有尖点。

在第二款，当单切面重点的平面和 F_3 的 3 条交线中只有一条实线时，F_3 沿 x 轴的左侧穿过二重切面。

综合起来，我们可以说：

为了考察清楚代数曲面二切面重点和单切面重点本身的形状，不仅要考虑二次项，而且还要考虑三次项。

§5 F_3 的形状概述

上面所谈，已经接近于能说明三次曲面的形状，因此，我要结合一般结论，作进一步考察。

我可以先给出 F_3 的理论的历史资料，首先是

1849 年，A·凯莱与 G·萨尔蒙，他们证明了在 F_3 上有 27 条直线（参看 *Cambridge and Dublin Math. Journal* 第 4 卷，第 118，252 页）。

那以后，研究了这些直线的组合关系，特别是

1863 年，L·施莱夫利具体讨论了什么时候这些直线是实的或虚的，还考察了 F_3 上可能有什么样的奇点（参看 *Phil. Transactions* 第 153 卷，1863 年，第 193—241 页）。

1873 年，上面已提到了，在 *Math. Ann.* 第 6 卷上，我本人关于 F_3 形状的研究。

与此相联系，有：

1879 年，C·罗登贝尔格（Rodenberg）在 *Math. Ann.* 第 14 卷，第46—116页，用解析方法验证了我通过几何连续性考虑所得的结果。另一方面，罗登贝尔格还密切结合我的研究，在 Verlage von Brill（现在是 Schilling）制造了一系列的模型。

现在，在我开始讨论 F_3 的形状时，我遵循的指导思想是，以第 257 页中已讨论过的具 4 个实二重点的图为出发点，通过连续过渡（化解其二重点）以得到较一般的曲面①。

在上述作为出发点的曲面上，我们已经可以验证 27 条直线的存在，因为那个四面体有 6 条棱，每条计算 4 次，此外还有 3 条水平面直线。

当我们以那个 F_3 为出发点时，一切从它经过直射变换所得到的曲面当然都要看做是等价的。不然的话，F_3 的多样性——它的方程含有 19 个常数——将使我们根本难以窥其全貌。在平面上，对椭圆、抛物线、双曲线不加区别，也正是基于这个观点。

理论上就是如此。实践上，还要具体地运用直射变换来说明这个观点。我们首先要了解经过直射变换后，那个初始曲面可能获得的不同形状。例如可以用各种平面和曲面相交，再分别把这些作为消失平面投射到无穷远处，考察这时曲面的形状，作成模型，然后找出它们的共性而把它们的区别放在一边，不予考虑。只有在了解它们的具体特点以后，才能对特点加以忽略。在那以前，绝不能那样做，于是有以下结论

正如我们按二次曲线同无穷远直线的关系来区别椭圆和双曲线（在这里，抛物线作为过渡款不列入）那样，我们可以把具 4 个实二重点的 F_3 按其和无穷远面的关系区别为 5 种类型，它们

---

① 可以和上面对于 n 次代数曲线的作法作类比：在那里，我们从一个具许多二重点的曲线（即已分解为低次曲线的曲线）出发，化解其二重点以得到其近邻曲线。

都已作成模型①。对这 5 种类型以后将不再明显地加以区别。

现在我们把这个曲面的二重点化解，以过渡到近邻曲面。

为此，我们联系到对平面图的作法。显然，我们可以把具一个二重点的曲线用两种方法化解其二重点，以得到无二重点的曲线(图 24 中的 1 和 2)。这两种方法并无本质区别。但是，在三维空间，如果把经过二重点的纵线旋转，就得到性质不同的曲面。一个是(1) 像双叶双曲面那样分成两块不相连的曲面；另一个是(2) 像单叶双曲面那样的曲面，即代替原先二重点附近的是当中束小了的一片曲面(图 25)。于是有定理：在过渡到近邻曲面中，曲面的一个二重点可以用两种方式化解：一种是分离；一种是融合。下面我们用记号“＋”表示融合，记号“－”表示分离。

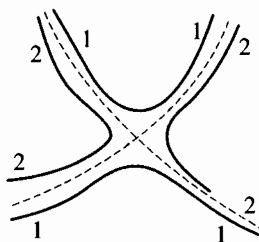


图 24

对初始曲面的每一个二重点都分别采取两种化解方式之一，以

① [这 5 种射影等价的具 4 个二重点的 F_3 可以从图 19 中的曲面按以下方法得到：直射变换的消失平面只和图 19 中的曲面的外面部分相交，而且

(1) 交于一个单支的 C_3(不含卵形线)，得到的曲面类型和图 19 相同；

(2) 交于分为含有两支的 C_3(有卵形线)。

消失平面和曲面中的四面体形部分相交，交线总是一个含有卵形线的 C_3。

(3) 4 个二重点都在消失平面的同一侧；

(4) 3 个二重点在消失平面的一侧，第四个在另一侧；

(5) 两个二重点在消失平面的一侧，其余两个在另一侧；

还会出现下列特款：

消失平面和曲面相切，且

(a) 切点在曲面的外面部分：交线是具普通二重点的 C_3；

(b) 切点在四面体形部分：交线是具孤立二重点的 C_3。

消失平面经过

(c) 曲面的一个二重点：交线是具普通或孤立二重点的 C_3；

(d) 曲面的两个二重点：交线分解为一条 C_2 和一条直线；

(e) 曲面的三个二重点：交线分解为 3 条直线；

(f) 消失平面含曲面上 3 条计算一次的直线；

(g) 消失平面含有一条连接曲面两个二重点的直线。]

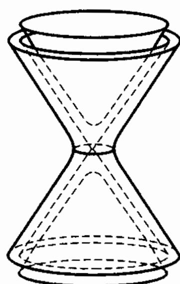


图 25

得到近邻曲面，这就有 5 种可能性，如表 3 所示。

表 3

Ⅰ.	＋＋＋＋
Ⅱ.	＋＋＋－
Ⅲ.	＋＋－－
Ⅳ.	＋－－－
Ⅴ.	－－－－

若考察这些曲面的形状，我们就会发现，在这 5 种无奇点的曲面中，有 4 种都是单块曲面，一种是双块曲面(－－－－)。这最后一种含一个卵形状部分和一个弯曲的，围绕它的部分。

这 5 款正好对应于施莱夫利在分析 F_3 上的直线的虚实时所划分的 5 种类型。

这个结论是通过如下的考虑得到的：任意取两个结点，化解时只要对其中至少一个采取了分离方式，连接这两个结点的直线就在所得曲面上化为 4 条虚直线。另一方面，若化解时，对两个结点都采取融合方式，则在过渡到近邻曲面时，连接结点的直线就化为 4 条实直线。按照这个规律，5 种曲面上的实直线数目，如表 4 所示。

表 4

Ⅰ.	$3+6\cdot 4=27$
Ⅱ.	$3+3\cdot 4=15$
Ⅲ.	$3+1\cdot 4=7$
Ⅳ.	$3+0\cdot 4=3$
Ⅴ.	$3+0\cdot 4=3$

其中最后一种和前面一种的区别是，它分成两块。这个结果和施莱夫利完全一致。

这里所得到的施莱夫利 5 种类型无奇点曲面同时又真实地代表

着拓扑学意义上的 5 种曲面：这就是说，属于同一个施莱夫利类型的两个曲面可以通过连续变形互相转化，当中不需要经过具二重点的曲面；或者说，若把三次曲面用 19 维空间的点来代表，则代表同一类型的点在该空间里构成连通域。

因此，从初始具 4 个二重点的曲面 F_3 通过＋和－的方法获得的曲面正好代表着施莱夫利 5 类曲面。当然，为了直观地说明，还必须对所得曲面通过直射变换加以转化，或者取具 4 个二重点的 5 类 $F_3$①，分别运用方法＋和－来转化。

但是，从我们那个具 4 个二重点的 F_3，通过连续变形，不但可以得到无二重点的 F_3，而且可以得到有任意奇点的 F_3。对此，罗登贝尔格(Rodenberg)的模型给出了许多的例子，可惜不能进一步介绍。

我们还要问：如何用解析方法②处理我们探讨中的曲面 F_3，它的最佳方程采取什么形式？

我指出 J·J·西尔维斯特的发现，即有唯一的一个所谓的五面形，它和曲面上的直线有密切联系。

为了研究这个曲面，选取这个五面形的 5 个平面为所谓的齐次五面形坐标的坐标平面。关于这种坐标，可说明如下。

取 5 个平面

$$x_i = a_i x + b_i y + c_i z + d_i = 0 \quad (i = 1, 2, 3, 4, 5)$$

所构成的五面形。它们中任意 4 个不经过同一点，任意 3 个不经过同一直线。对于空间任意点，取由这 5 个平面到它的距离，乘以一定的常数，作为它的坐标，即用这样所得的一组数 $x_1 : x_2 : x_3 : x_4 : x_5$ 表示该点。由于 4 个平面已足以确定点的齐次坐标，这 5 个式 x_i 不能相互独立。在超过需要数的这 5 个 x_i 之间，有一个齐次线性关

① 这是指第 262 页所说的，按 F_3 和无穷远面的关系的分类。——中译者

② [关于用解析方法处理三次曲面上直线的问题，新近有范德瓦尔登(B. L. v. d. Waerden)的一篇文章："Der Multiplizitätsbegriff der algebraischen Geometrie"("代数几何中的阶数概念")，*Math. Ann.* 第 97 卷，1927 年，第 756—774 页。]

系。设选取确定坐标系时那 5 个常数,使

$$x_1+x_2+x_3+x_4+x_5=0。$$

西尔维斯特发现,在这些他称之为五面形坐标系中,有一个最优的。在这个坐标下,曲面 F_3 的方程只出现 x_i 的三次方。于是曲面方程就成为

$$\sum_{i=1}^{5} a_1 x_i^3 \quad (其中\sum_{i=1}^{5} x_i=0)。$$

罗登贝尔格利用西尔维斯特这个标准方程研究了所描写的几何关系。标准方程只含有曲面方程中 19 个常数中的 4 个,而这 4 个是必须留下的。因为通过适当的坐标变换,只能去掉其中 15 个。在这里,平面 $x_i=0$ 当然是可实也可虚的。

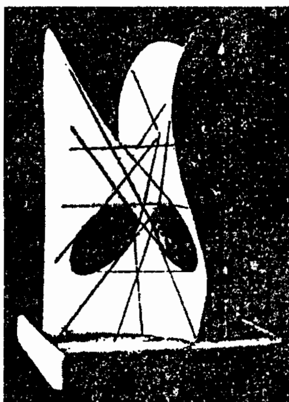


图 26

特殊地,我们专考虑实的五面形,并且从西尔维斯特方程得到一个特别简单的,就是使其中 a_i 相等的 F_3。这个曲面被克莱布什称为对角曲面。它有 27 条实直线,这曲面准确地在图 26 中表示着。这个模型是在 A · 克莱布什的鼓励下由威勒(A. Weiler)制作的,它对我关于三次曲面的研究提供了帮助。

这个曲面有什么特殊性? 何以它有那个名称?

令五面形坐标的平面 $x_1=0$ 和其他坐标平面以及曲面相截,就得到四边形 $x_2=x_3=x_4=x_5=0$ 以及曲面的截痕

$$x_2^3+x_3^3+x_4^3+x_5^3=0。$$

现在利用关系 $x_2+x_3+x_4+x_5=0$,可以证明,曲面的截线为上述

四边形的 3 条对角线所构成。这就是曲面名称的来源，因此，"对角曲面"的名称来自这样的事实：曲面和五面形的每个平面交于该平面和其余 4 个平面的交线所构成的四边形的 3 条对角线。

于是已得到曲面上 (3 · 5 =)15 条直线，关于其余 12 条直线，我在此不进一步说明；你们不难在模型上找到①。

关于那 15 条直线还有一个定理：它们 3 条一组，交于 10 点。我们只需想一想，经过任意五面形的每个顶点，即 3 个面的交点，有 3 条(四边形的)对角线，分别在 3 个平面上，而五面形有 10 个顶点。因此(简单地说)，15 条直线有 10 个交点，经过每个交点都有 3 条。这些交点是五面形的顶点。

就谈到此为止，你们对 F_3 的形状已了解其概要，至于细节，就

---

① [其余 12 条直线可以通过解析方法推得：设 1, ε, ε^2, ε^3, ε^4 为单位 1 的 5 个五次根。容易看出

$$1+\varepsilon+\varepsilon^2+\varepsilon^3+\varepsilon^4=0, \tag{1}$$

$$1^3+\varepsilon^3+(\varepsilon^2)^3+(\varepsilon^3)^3+(\varepsilon^4)^3=0。 \tag{2}$$

由此可见，$1:\varepsilon:\varepsilon^2:\varepsilon^3:\varepsilon^4$ 代表对角曲面上一点，由于 1, ε, ε^2, ε^3, ε^4 这 5 个数的次序可以任意颠倒，它们仍然满足方程(1)和方程(2)，就一共有(5! =)120 组数值，它们都代表对角曲面的点。但在这 120 组数值中，彼此只差一个常数因子的都代表同一点。例如分别用 1, ε, ε^2, ε^3, ε^4 依次乘 1, ε, ε^2, ε^3, ε^4，考虑到 $\varepsilon^5=1$, $\varepsilon^6=\varepsilon$, $\varepsilon^7=\varepsilon^2$, $\varepsilon^8=\varepsilon^3$，就得代表同一点的 5 组数值

$$\begin{aligned} x_1:x_2:x_3:x_4:x_5 &= 1:\varepsilon:\varepsilon^2:\varepsilon^3:\varepsilon^4 \\ &=\varepsilon:\varepsilon^2:\varepsilon^3:\varepsilon^4:1 \\ &=\varepsilon^2:\varepsilon^3:\varepsilon^4:1:\varepsilon \\ &=\varepsilon^3:\varepsilon^4:1:\varepsilon:\varepsilon^2 \\ &=\varepsilon^4:1:\varepsilon:\varepsilon^2:\varepsilon^3。\end{aligned}$$

所以，那 120 个数值组只代表对角曲面上(4! =)24 个不同点。显然，若固定 $1:\varepsilon:\varepsilon^2:\varepsilon^3:\varepsilon^4$ 中的第一数，而把其余 4 个数作一切可能排列，就得到这 24 个点。这 24 个点包含 12 对共轭虚点。现在，把这 12 对共轭虚点逐对相连，所得的 12 条实直线都整条在曲面上。例如，若 $x_1:x_2:x_3:x_4:x_5$ 为 24 组数值之一组，而 $X_1:X_2:X_3:X_4:X_5$ 是它的共轭虚点，则在它们连线上的点是

$$(\lambda x_1+\mu X_1):(\lambda x_2+\mu X_2):\cdots:(\lambda x_5+\mu X_5)。$$

容易验证，对于任意 λ, μ 值，这组值满足方程(1)和方程(2)。]

要你们自行钻研了[①]。

呼吁：通过观察自然，不断修订传统科学结论

现在，请让我用下面的论点来结束我的讲演。

在这个讲演里，我讨论了通常在教科书里找不到的许多东西，但又无形中假定了你们了解教科书里的一般内容。我愿意通过这种方式促使你们自由思考，独立地作出判断，来掌握所讨论的事物。例如，请你们思考一下我对于经验曲线和曲面的说法，思考一下我的讨论一般地又限于解析的图形。

数学的情况犹如造型艺术。向先贤们学习不但有益，而且很有必要。但是，如果局限于学习传统的东西，而只是由书本学到的继续前进，就会产生我所谓的学院式的体系。与此相对立，我提出劝告：

保持一流大师的遗风：回到固有的生动活泼的思考，回到自然！

---

① [D·希尔伯特在他的论文："Über die Gleichung neunten Grades"(关于第九条直线的方程)，*Math. Ann.* 第 97 卷，1927 年，第 243—250 页中对 F_3 上含有 27 条直线的事实有优美的应用。]

人名译名对照

A

Abel，N.　　阿贝尔
Abraham，M.　　阿伯拉罕
Alexandroff，P.　　亚历山大罗夫
Ambronn，L.　　阿姆布朗
Ambronn，R.　　阿姆布朗
Amsler　　安姆斯拉
Apollonius　　阿波罗尼斯
Archimedes　　阿基米德
Aristoteles　　亚里士多德
Auwers　　奥沃斯

B

Bachman　　巴赫曼
Ball，B. St.　　波尔
Ball，R.　　波尔
Bauer，G.　　鲍尔
Baumann　　鲍曼
Baumeister，V.　　鲍梅斯特
Bauschinger，J.　　鲍申格尔
Behrendsen　　贝伦德逊
Beneke　　本内克

Berkeley	贝克莱
Bernoulli, D.	伯努利
Bernoulli, J.	伯努利
Bernoulli	伯努利
Bernstein, F.	伯恩斯坦
Bertrand, J.	贝特兰
Bessel, F. W.	贝塞尔
Bézout	贝佐
Bieber bach, L.	比贝尔巴赫
Blackburn, H.	布莱克本
Blaschke, W.	布拉施克
Boltzmann, L.	波尔兹曼
Bolya., J.	鲍耶
Bolyai, W.	鲍耶
Bolzano, B.	波尔查诺
Borel, E.	波雷尔
Boussintsq, J.	布森纳斯克
Brandl, Al.	布兰德尔
Brianchon	布赖恩昌
Briggs, H.	布里格斯
Brill, A.	布里尔
Brouwer, L. E. J.	布劳威尔
Brunn, H.	布龙
Bruns, H.	布龙斯
Burkhardt, H.	伯克哈特
Byerly, W. E.	拜尔利
Böcher, M.	博克
Böhmer	伯默尔

Börsch，A.	伯尔施
Bürgi，J.	比尔吉
v. Braunmühl，A.	布劳恩米尔

C

Cajori，F.	凯乔里
Cantor，G.	康托
Caradi	卡拉第
Carathéodory，C.	卡拉西奥多里
Cardan，G.	卡丹
Cardano	卡丹诺
Carleman，T.	卡尔曼
Cartesius	卡德修斯
Cauchy，A.	柯西
Cavalieri，B.	卡瓦列里
Cayley，A.	凯莱
Chernac	切纳克
Chisholm，G.	奇泽姆
Chrystal	克里斯特尔
Clebsch，A.	克莱布什
Clifford，W. K.	克利福德
Coble，A. B.	科布尔
Copernicus，N.	哥白尼
Coradi	科拉迪
Courant，R.	柯朗
Cremona	克里蒙那

D

Damaskics	大马斯克斯

Feigl, G.	费格尔
Fejér, L.	费耶尔
Fermat	费马
Feuerbach	费尔巴哈
Fiedler, W.	费德勒
Fine, H.	法因
Fine, H. B.	芬恩
Finsterwalder	芬斯特瓦尔德
Fleck, A.	弗莱克
Fleischer, H.	弗莱谢尔
Fourier	傅里叶
Fricke, R.	弗里克
Friesecke, H.	弗里泽克
Friesendorff	弗里森多尔夫
Föppl, A.	福普尔

G

Gauss, C. F.	高斯
Geiringer	盖林格
Gerhardt, K. J.	格哈特
Gergonne	格根
Geuer, F.	戈伊尔
Gibbs, J. W.	吉布斯
Gorden, P.	戈登
Gordon, P.	哥尔丹
Goursat, E.	古尔萨
Grassmann, H.	格拉斯曼
Groeneveld, J.	格勒内费尔德

Hesse, L. O.	黑塞
Heun	霍伊恩
Hilbert, D.	希尔伯特
Hilb	希尔伯
Hjelmslve, J.	耶尔姆斯勒夫
Hobson, E. W.	霍布森
Hollcroft, T. R.	霍考夫特
Homann, H.	霍曼
Hölder, O.	霍尔德
Huntington	亨丁顿
Hurewicz, W.	胡维尔兹
Hypsikles	海珀西斯

I

Itzigsohn	伊特齐格桑

J

Jacobi, G. J.	雅可比
Jordan, C.	若当
Jordan, W.	若当
Juel, C.	于尔

K

Kant, I.	康德
Kempe, A. B.	肯普
Kepler, J.	开普勒
Kerékjartó, B. V.	克里克贾多
Kimura	木村

Kirsch	基尔施
Klein, F.	克莱因
Kneser, A.	克内泽尔
Knidos	克尼多斯
Knight	乃特
Knopp. K.	克诺谱
Koemgsberger, L.	柯尼希斯伯格
Kommerell, K.	科默莱尔
Kowalewski, G.	柯瓦列夫斯基
Kronecker, L.	克罗内克
Kummer, E.	库默尔
Kästner, A. G.	卡斯特纳
Köpke, A.	科普克
König	康尼格

L

Lacroix	拉克罗依克斯
Lagrange, L.	拉格朗日
Lamé	拉梅
Langel, L.	兰格尔
Laplace, P. S.	拉普拉斯
Legendre	勒让德
Leibniz, G.	莱布尼兹
Lemoine	莱莫恩
Lie, S.	李
Lietzmann, W.	里茨曼
Lilienthal, R.	利林塔尔
Lindermann, F.	林德曼

Liouville, J.	刘维尔
Lobatschefsky	罗巴切夫斯基
Lohmann, W.	洛曼
Lorey, W.	洛莱
Lotlermoser	洛特莫泽
Ludwig, W.	路德维格
L'Hospital	洛必达
Lübson	卢勃森

M

Maclaurin, C.	麦克劳林
Maennchen, Ph.	曼恩钦
Mangoldt, V.	曼戈尔特
Markoffs	马尔柯夫斯
Markoff	马尔科夫
Maser	马塞
Maxwell, J.	麦克斯韦
Meder, A.	梅德尔
Mehmke	密姆克
Menger, K.	门格尔
Mercator, N.	默卡多
Meyer, W. Fr.	梅耶
Michelson, A. A.	迈克森
Minkowski, H.	闵可夫斯基
Molk, J.	莫尔克
Mollweide	摩尔维迪
Morrice, G. C.	莫里斯
Moser	莫泽

Pascal, Bl.	帕斯卡
Pasch, M.	帕什
Peano, G.	皮亚诺
Peaucellier	波塞利
Perrin, J.	佩林
Perron, O.	佩龙
Perry, J.	佩雷
Peurbach, G.	普尔巴赫
Piard	皮卡
Pitiscus	皮蒂斯克斯
Pizzetti, P.	皮策蒂
Plato	柏拉图
Plücker, J.	普吕克
Pohlke, K.	波尔克
Poincaré, H.	庞加莱
Poinsot	普昂索
Poisson, S. D.	泊松
Pollack, L. W.	波拉克
Pollazek, H.	波拉泽克
Poncelet, J. V.	彭赛列
Pothenot	波特诺特
Pressler, O.	普雷斯勒
Pringsheim, A.	普林斯海姆
Proclus, D.	普罗克拉斯
Prümm	普卢姆
Ptolemy	托勒密
Pythagoras	毕达哥拉斯

R

Regiomontanus	雷乔蒙努司
Reinhardt，C.	莱因哈特
Remak，R.	雷马克
Rhäticus	软蒂克斯
Riemann	黎曼
Riesz，M.	黎斯
Riesz	里茨
Ritter，E.	利特
Robiuson，G.	罗宾森
Rodenberg，C.	罗登贝尔格
Rolle	罗尔
Rosenbaum，W.	罗森鲍姆
Rudio	鲁迪奥
Runge，C.	龙格
Rósenthal，A.	罗森塔尔

S

Salmon，G.	萨尔蒙
Schafheitlin，P.	谢弗里特林
Scheffers，G.	谢弗思
Scheffer	谢弗
Schellbach	谢尔巴赫
Schepp，A.	舍普
Schilling	席林
Schimmack，R.	席马克
Schlömilch	施洛米尔赫

Schläfli, L.	施莱夫利
Schmidt, Ad.	施密特
Schoenflies, A.	舍恩弗列斯
Schubert, H.	舒伯特
Schuh, Fr.	舒
Schwartz, H. A.	许瓦兹
Schweikart	施韦卡特
Schwerdt, H.	施韦尔特
Serret	塞雷特
Seyfarth	赛法特
Shanks, W.	向克斯
Simon, M.	西蒙
Simpson, Th.	辛普生
Snellius	斯涅尼奥斯
Sonin	宗宁
Stäckel, P.	施塔克尔
Staut, V.	施陶特
Steffersen, J. F.	斯蒂芬森
Steiner, J.	斯坦纳
Stifel, M.	史蒂菲尔
Stolz, O.	斯托兹
Study	斯特迪
Sturm	斯图姆
Süchting, Fr.	聚希廷
Sylvester, J. J.	西尔维斯特
v. Sanden, H.	桑登
v. Sanden, H.	赞登

T

Tannery, J.	特纳里
Taylor, B.	泰勒
Thomson, W.	汤姆逊
Tietze, H.	蒂采
Timerding, E. E.	蒂默丁
Tissot	蒂索
Tropfke	特罗夫克
Tsecheby Seheff, P. L.	切比雪夫
Tägert, F.	塔格特

U

Uryson, P.	乌里松

V

Vahlen Th.	范伦
Varignon	瓦里农
Vega	维加
Vermeils, H.	弗尔迈尔斯
Vermeil	弗尔迈尔
Vermeit, H.	弗尔梅特
Veronese, G.	韦龙尼斯
Vieta	韦达
Vlacq, A.	弗拉克
Voigt, W.	伏格特

W

v. d. Waerden, B. L.	范德瓦尔登
Walther, A.	瓦尔特
Weber, H.	韦伯
Weierstrass, K.	魏尔斯特拉斯
Weiler, A.	威勒
Wellstein, J.	韦尔斯坦因
Wendland, P.	温德兰
Weyl, H.	外尔
Whittaker, E. T.	惠特克
Wibraham, H.	威尔布里厄姆
Wiener, Chr.	维纳
Wiener, N.	维纳
Wiles, A.	怀尔斯
Wilson, E. B.	威尔逊
Wittenberg	魏登堡
Witting	维廷
Wolff, C.	沃尔夫
Wolfskehl	华尔夫斯凯尔
Wüllner	沃尔纳

Y

Young, J. W. A.	杨

Z

Zeuthen	佐登
Zimmer mann, H.	齐末尔曼

Zipperer，L.	齐佩雷尔
Zühlke，P.	齐尔克
Zülke，P.	左尔克
Фихтеиропьд，Г. М.	菲赫金哥尔茨

译后记

1965年前后，高等教育出版社经广泛试稿后，决定请我的老师陆秀丽教授翻译19世纪末、20世纪初德国知名数学大师菲尼克斯·克莱因所著的《高观点下的初等数学》(以下简称《高观点下》)一书。1966年史无前例的"文化大革命"运动到来，陆老师的翻译工作不仅中断，而且不得不将译出的部分书稿付之一炬。20世纪80年代初，上海科技出版社登门邀请陆老师再译《高观点下》一书，陆老师不愿旧事重提，婉拒了上海科技出版社的善意约请。

20世纪80年代初期，我开始阅读《高观点下》(英文版)，深深感到这里的初等数学内容并不初等，书中文字也不是普通的数学语言，很少运用数学形式推导；它所讨论的并非一般的数学教材内容，书的写作结构也别具一格，不同于过去我读过的任何一本数学书籍。克莱因在书中对各个分支的数学概念、语言、符号以及运算法则等各个环节中问题的产生与发展，其内因、外因和发展过程的思维描述都非常细腻、清晰、灵活，且逻辑性很强，其中还涉及接受数学知识的心理学和认知论问题的讨论。此书所涵盖的知识广泛，包括哲学、物理学、天文学、气象、测量等多个学科中的许多问题，有些问题十分复杂。众所周知，自17世纪以来，数以百计的世界级大科学家如牛顿、康托、伽利略、开普勒、莱布尼兹、庞卡莱等人，对数学领域方方面面问题的提出、论证以及应用都曾取得重大成就，克莱因将其融入数学这一严谨学科的各个分支的各个部分，揭示它们的共性、它们的本质以及它们之间的相互关系，使《高观点下》一书成为完整和谐的有机整体。尤其是书中始终关注数学教育，强调数学教学的方向、教师必须注重的数学知识以及有效的教学方法，旨在提高教师的一般数学素养。阅读《高观点下》这一丰富多彩、内涵极其深刻的巨著是有一定难

度的，但凡是具有一定数学知识的人都可以从中获得教益和启发。

在我阅读学习的过程中体会到《高观点下》一书的意义非同寻常，我决心将它翻译出来，争取出版，以供数学科学工作者及教师阅读、参考。

我和陈义章、杨钦樑 3 人对《高观点下》一书分译互校、反复讨论修改，在两年的时间里终于艰苦地将《高观点下》第一、第二卷译完。为了保持这一世界巨著原有的学术水准，保证翻译质量，我们请时任武汉大学校长的齐民友教授审阅《高观点下》第一卷，请时任武汉大学中法文化交流中心主任的余家荣教授审阅《高观点下》第二卷。两位教授在百忙之中对译稿逐章逐节、逐字逐句校正修改，他们的审阅使译文中很多概念表述更为准确清晰，论证更为严谨完美，对提高全书翻译质量起到了重要的作用。

为了引起广大读者的注意，同时也是为读者提供有益的阅读参考，特请知名学者吴大任教授撰文介绍、推荐《高观点下》一书。吴老在较短的时间里写出了“博洽内容　独特风格——介绍克莱因《高观点下的初等数学》”一文交由《数学通报》(1989 年第 6 期)发表。此外，吴老还为《高观点下的初等数学》中译本初版写了序言。

《高观点下的初等数学》第一、第二卷中译本得以在 1989 年由湖北教育出版社出版，应当说正是数学界前辈们无私帮助、大力支持的结果。在对《高观点下的初等数学》一书的审阅、推荐等过程中，我深切体会到数学界前辈们对数学科学的无比热爱，对世界名著《高观点下的初等数学》一书的珍视，以及他们对学术事业无私的奉献精神和责任感，这种崇高的风范值得我们永远学习。

值此《高观点下的初等数学》中译本新版出版之际，特向帮助和支持过我们的齐民友教授、余家荣教授和陆秀丽教授诚恳致谢，特别缅怀吴大任教授。

对本书译文中不妥之处，请读者批评指正。

舒湘芹

2007 年 8 月

lementary Mathematics from an Advanced Standpoin

教师应该具备更高的数学观点。理由是，观点越高，事物越显得简单。

《高观点下的初等数学》一书，至今读来仍然感到十分亲切。这是因为，其内容主要是基础数学，其对
点蕴含着真理，而当时德国数学教育中的不少问题，在今日之我国也仍然存在。克莱因声称本书是为中学教
师和成熟的大学生写的，但按其内容，所有对数学有一定了解的人都可以从中获得教益和启发……现代数学
已发生了极大变化，新成果、新概念、新观点、新学科层出不穷。我热切希望我国高水平的数学多面手会写
出更结合我国实际的、现代化的《高观点下的初等数学》。这样一本书的出版将是我国数学教育史上的一件
大事。

■ 吴大任

读这本书，您会感到极有收获，而不得不心悦诚服。不得不承认克莱因是真正的大师！

■ 齐民友

除了数学的工作之外，克莱因的数学史至今仍是19世纪数学史上的重要的标准著作，作为当时的领袖数
学家，他的许多观点至今仍然对数学家、数学史家有所启迪。他的《高观点下的初等数学》反映了他对数学
的许多观点，是一本译为多种文字的通俗读物，影响至今不衰。

■ 胡作玄

菲利克斯·克莱因教授是德国有名的数学研究家，他也是一位循循善诱的教师。他以罕见的天才，集一
切数学领域的知识于一身，并善于领悟这一切领域之间的相互关系。他认为使学生了解数学并不是孤立的各
门学问，而是一个有机的整体，是他作为一个教师的明显职责。他对中学数学教学有浓厚的兴趣，不仅关心
应该教些什么内容，而且关心怎样教才是最有效的方法……他一贯努力缩短中学和大学之间的差距，从传统
的漠不关心中激起中学教师对高等数学的兴趣，把中学数学教学引向健康发展的方向；同时也努力扭转大学
的态度及教学方向，使之承认中学的正常地位，使数学教育前后一贯……《高观点下的初等数学》是一本无
比珍贵的著作，同样可作为大学教师和中学教师的参考书。无论就材料安排的巧妙或就讨论方式的引人入胜
来说，目前都没有一本书可以同这本书相比。

■ 洛杉矶加利福尼亚大学数学教授　E·R·赫德里克
■ 伯克利加利福尼亚大学数学教授　C·A·诺布尔

ISBN 978-7-309-05982-3



责任编辑：范仁梅
美术编辑：陈　萍

定价：68.00元